Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Science

presented by

Anne Rademacher, M.Sc.

born in Wilhelm-Pieck-Stadt Guben, Germany

Oral examination: November 28, 2017

Silencing and activating heterochromatin

Referees: PD Dr. Karsten Rippe
appl. Prof. Dr. Harald Herrmann-Lerdon

On ne fait jamais attention à ce qui a été fait;
on ne voit que ce qui reste à faire.

Marie Skłodowska Curie

Abstract

Repetitive, gene-poor regions termed heterochromatin are typically transcriptionally silent and mostly located at specific chromosomal loci such as (peri)centromeres and telomeres, but are also found interspersed throughout the rest of the genome. Heterochromatin silencing is associated with epigenetic features such as DNA methylation and post-translational modification of histone tails, e. g. trimethylation of histone 3 lysine 9 (H3K9me3), which is set by the histone methyltransferases SUV39H1 and SUV39H2 among others. However, the mechanisms underlying establishment and stability of repressive histone marks as well as their contribution to transcriptional silencing remain unclear. Here, I show that the SUV39H histone lysine methyltransferases recognize heterochromatin downstream of the methyl-DNA binding protein MECP2 and remain chromatin-bound throughout the cell cycle, linking these enzymes to the specificity and memory of heterochromatin. In addition, SUV39H isoforms displayed distinct non-redundant functions. SUV39H2 contributed more to establishing H3K9me3 than SUV39H1 in mouse embryonic fibroblasts and loss of SUV39H-dependent H3K9me3 was not sufficient for transcriptional derepression. Since tools to resolve the kinetics of transcriptional activation with high spatial and temporal resolution are currently scarce, I devised the Blue Light-Induced Chromatin Recruitment (BLInCR) tool to dissect the activation of the repressed heterochromatin state. I used BLInCR to rapidly and reversibly target the viral transactivator VP16 to an integrated reporter in the U2OS 2-6-3 cell line and to measure the activation kinetics of an initially repressed reporter gene locus. First transcripts were detectable within minutes after triggering the relocalization of VP16 and activation progressed in two phases. Thus, even a strong activator was not able to activate silenced gene clusters in one step. The BLInCR tool introduced here enables the use of complex activation patterns, which will also be valuable to assess the effect of repeated activation and, more generally, to probe the stability of chromatin states to understand the regulation of epigenetic memory. Taken together, these results indicate that only some genes are susceptible to SUV39H- and H3K9me3-mediated repression and that the heterochromatin state can be overwritten by strong, targeted activators. The insights gained within this study on the mechanisms of silencing and activating heterochromatin can help understand heterochromatic deregulation. The latter is frequently observed in disease, e. g. in cancer, and thus, the heterochromatin network might harbor promising targets for epigenetic drug development.

Zusammenfassung

Repetitive Teile des Genoms, welche wenige Gene enthalten und in der Regel nicht transkribiert werden, werden als Heterochromatin bezeichnet. Diese Sequenzen befinden sich überwiegend in speziellen Regionen der Chromosomen wie den (Peri-)Zentromeren oder Telomeren, sind aber auch im übrigen Genom vorhanden. Die transkriptionelle Inaktivität des Heterochromatins wird durch epigenetische Faktoren wie DNA-Methylierung und posttranslationale Modifikationen von Histonproteinen gewährleistet. Eine solche Modifikation ist die Trimethylierung von Lysin 9 des Histons 3 (H3K9me3), welche unter anderem durch die Histonmethyltransferasen SUV39H1 und SUV39H2 katalysiert wird. Durch welche Mechanismen repressive Histonmodifikationen etabliert und stabilisiert werden und welche Rolle sie in der transkriptionellen Inaktivierung spielen, ist bisher weitestgehend unklar. In meiner Arbeit zeige ich, dass die Erkennung von Heterochromatinregionen durch SUV39H-Enzyme dem methyl-DNA-bindenden Protein MECP2 untergeordnet ist. Außerdem bleiben sie während des Zellzyklus‘ an Chromatin gebunden und könnten somit zur Spezifität und Erhaltung von Heterochromatin beitragen. Des Weiteren haben die beiden SUV39H-Isoformen unterschiedliche Funktionen. SUV39H2 trägt in embryonalen Mausfibroblasten mehr zur Trimethylierung von H3K9 bei als SUV39H1 und der Verlust von SUV39H-abhängiger Trimethylierung von H3K9 ist nicht hinreichend für eine transkriptionelle Reaktivierung. Da es derzeit nur wenige Strategien zur Analyse der Kinetik der Transkriptionsaktivierung mit hoher zeitlicher und räumlicher Auflösung gibt, habe ich ein Verfahren zur lichtinduzierten Rekrutierung von Effektoren an Chromatin namens BLInCR entwickelt, um die Aktivierung eines inaktivierten Lokus zu untersuchen. Mit BLInCR konnte der virale Aktivator VP16 schnell und reversibel an einen heterochromatischen Reporterlokus in der U2OS 2-6-3-Zelllinie gebunden werden um dessen Aktivierungskinetik zu messen. Die Aktivierung konnte innerhalb weniger Minuten nach der Bindung von VP16 durch RNA-Visualisierung detektiert werden und lief in zwei Phasen ab. Das bedeutet, dass selbst ein starker Aktivator einen inaktivierten Lokus nicht in einem Schritt anschalten konnte. Das hier vorgestellte BLInCR-Verfahren ermöglicht die Anwendung verschiedenster Beleuchtungsmuster, was genutzt werden kann um mehrfach Aktivierung zu initiieren. Im Allgemeinen kann so die Stabilität von Chromatinzuständen geprüft und die Erhaltung von epigenetischen Zuständen untersucht werden. Die hier präsentierten Ergebnisse zeigen, dass nur einige Gene SUV39H- und H3K9me3-vermittelt inaktiviert werden und dass der heterochromatische Zustand durch starke Aktivatoren überschrieben werden kann. Die Erkenntnisse aus dieser Arbeit hinsichtlich der Mechanismen, welche die transkriptionelle Inaktivierung und Reaktivierung von Heterochromatin steuern, können hilfreich sein, um ihre Deregulation zu verstehen. Diese tritt häufig im Krankheitsfall, z. B. bei Krebs, auf und somit könnten im hier untersuchten Heterochromatinnetzwerk Kandidaten zur Entwicklung epigenetischer Medikamente zu finden sein.

List of publications

During this thesis, I contributed to the following publications:

K. Muller-Ott, F. Erdel, A. Matveeva, J.P. Mallm, A. Rademacher, M. Hahn, C. Bauer, Q. Zhang, S. Kaltofen, G. Schotta, T. Hofer & K. Rippe. 2014. Specificity, propagation, and memory of pericentric heterochromatin. *Mol. Syst. Biol.* 10:746.

A. Rademacher, F. Erdel, J. Trojanowski, S. Schumacher & K. Rippe. 2017. Real-time observation of light-controlled transcription in living cells. *bioRxiv* doi: 10.1101/132050. Under review at the *Journal of Cell Science*.

Contents

List of Figures

Abbreviations

5-aza-C	5-aza-2'deoxycytidine
5mC	5-methyl-cytosine
*lac*O	*lac* operon
*tet*O	*tet* operon
APD	avalanche photodiode
BLInCR	Blue Light-Induced Chromatin Recruitment
BRD	bromodomain-containing protein
BSA	bovine serum albumin
CD	chromo domain
cDNA	complementary DNA
ChIP	chromatin immunoprecipitation
ChIP-seq	chromatin immunoprecipitation followed by DNA sequencing
CIB1	cryptochrome-interacting basic helix-loop-helix 1
CIBN	N-terminal domain of CIB1
CLSM	confocal laser scanning microscopy
CpG	cytosine-phosphate-guanine
CRY2	cryptochrome 2
CTD	C-terminal domain
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA-binding domains
DMEM	Dulbecco's Modified Eagle Medium
dn	double null
dNTP	deoxynucleoside triphosphate
DSIF	DRB-sensitivity-inducting factor
DTT	Dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ERV	endogenous retrovirus
FBS	fetal bovine serum
FISH	fluorescence *in situ* hybridization
FP	fluorescent protein
FRAP	fluorescence recovery after photobleaching
FWHM	full width at half maximum
gnls	non-linear generalized least squares method
GSAT	gamma satellite repeats
H2Bub	monoubiquitinated histone 2B
H3K14ac	histone 3 lysine 14 acetylation
H3K27me3	histone 3 lysine 27 trimethylation
H3K36me3	histone 3 lysine 36 trimethylation
H3K4me3	histone 3 lysine 4 trimethylation
H3K9ac	histone 3 lysine 9 acetylation
H3K9me3	histone 3 lysine 9 trimethylation
H4K20me3	histone 4 lysine 20 trimethylation
hADD	human ATRX-Dnmt3-Dnmt3L domain of the ATRX protein
HD	homeodomain
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	human immunodeficiency virus
HP1	heterochromatin protein 1
HRP	horseradish peroxidase
IF	immunofluorescence
IGV	Integrative Genomics Viewer
iMEF	immortilized mouse embryonic fibroblast
KAT	histone acetyl transferase

KMT ..... lysine methyl transferase
LAD ..... lamina-associated domain
LINE ..... long interspersed nucleotide element
LOV ..... light-oxygen-voltage
LTR ..... long terminal repeat
MBD ..... 5-methyl-cytosine binding domain
mBRD2 ..... murine bromodomain containing protein 2
mCD ..... chromo domain of mouse HP1β
MCP ..... MS2 coat protein
MEF ..... mouse embryonic fibroblast
mESC ..... mouse embryonic stem cell
NELF ..... negative elongation factor
NLS ..... nuclear localization signal
NTP ..... nucleotide triphosphate
P-TEFb ..... positive transcription elongation factor-b
PBS ..... phosphate buffered saline
PCH ..... pericentric heterochromatin
PCR ..... polymerase chain reaction
PD ..... paired domain
PEV ..... position effect variegation
PFA ..... paraformaldehyde
PHYB ..... phytochrome B
PIC ..... pre-initiation complex
PIPES ..... piperazine-N,N'-bis(2-ethanesulfonic acid
PMSF ..... phenylmethylsulfonyl fluoride
pol II ..... DNA-directed RNA polymerase II
PTM ..... post-translational modification
qRT-PCR ..... quantitative real-time PCR
Rb ..... retinoblastoma
RNA-seq ..... RNA sequencing
ROI ..... region of interest
RPKM ..... reads per kilobase of exon per million mapped reads
RT ..... room temperature
SDS ..... sodium dodecyl sulfate
sem ..... standard error of the mean
SETDB ..... SET domain bifurcated
sn ..... single null
SSC ..... saline-sodium citrate
TAD ..... transactivation domain
TAE ..... Tris base, acetic acid, EDTA
TAF ..... TBP-associated factor
TALE ..... transcription activator-like effector
TBME ..... *tert*-butyl methyl ether
TBP ..... TATA box-binding protein
TBS ..... Tris buffered saline
TCA ..... trichloroacetic acid
TE ..... Tris EDTA
TF ..... transcription factor
TSS ..... transcription start site
VRC ..... vanadyl ribonucleoside complex
WB ..... western blot
wt ..... wild type
Zfp ..... Zinc finger protein

I Introduction

Eukaryotic genomes are efficiently packaged into nuclei through interactions with a variety of proteins and RNA – an assembly termed chromatin. Traditionally, loosely packaged, gene-rich euchromatin is distinguished from tightly packaged, gene-poor heterochromatin (Heitz 1928). The latter comes in two flavors: the invariantly condensed constitutive heterochromatin and the cell type-dependent facultative heterochromatin. In general, constitutive heterochromatin contains a high proportion of repetitive sequences, is replicated late in S-phase and lacks recombination. In contrast, facultative heterochromatin denotes genomic regions that are only 'heterochromatinized' in a subset of cells or only for one homolog (Richards & Elgin 2002). Importantly, even though heterochromatin appears denser, it is not dense enough to virtually exclude proteins due to its spatial configuration thereby preventing DNA binding. Reportedly, proteins $\leq$10 nm in diameter (corresponding to $\sim$1 MDa) display basically no accessibility differences in hetero- compared to euchromatin (Gorisch *et al.* 2005, van Steensel 2011). Differences in transcription between those two major compartments are therefore generally attributed to post-translational modifications (PTMs) of histone tails, DNA methylation or differential enrichment of other protein factors that regulate the accessibility of DNA on the molecular level (van Steensel 2011). Due to the variety of features that have been described so far, efforts have been made to find functional combinations and segregate all possible modification patterns into a defined set of chromatin states (Baker 2011). Diverse cell systems from plant seedlings (Roudier *et al.* 2011) to human lymphocytes (Ernst & Kellis 2010) have been analyzed by diverse statistical approaches and yielded different numbers of chromatin states. Those states are associated with different functions including repression of developmental genes in differentiated cells, silencing of repetitive sequences or active transcription (van Steensel 2011).

I.1 Pericentric heterochromatin – a model system for constitutive heterochromatin

Pericentric heterochromatin (PCH) refers to the centromere-flanking regions that form a distinct chromatin compartment that is associated with repressed chromatin states. The containing repeat sequences and their organization in the nucleus vary greatly between different organisms suggesting that epigenetic regulation rather than binding of conserved DNA sequence motifs govern PCH formation (Saksouk *et al.* 2015). In mice, the adjacent centromers harbor the 120 bp minor satellite repeats (Pietras *et al.* 1983, Wong & Rattner 1988), whereas PCH comprises the AT-rich, 234 bp major satellite repeat DNA (Horz & Altenburger 1981, Manuelidis 1981) (**Fig. 1a**). Murine PCH makes up $\sim$3% of the mouse genome (Martens *et al.* 2005) and is a model system for constitutive heterochromatin

since it forms distinct nuclear patterns in interphase nuclei of differentiated cells. In mouse embryonic fibroblasts (MEFs) PCH can be easily distinguished from surrounding euchromatic regions by light microscopy (**Fig. 1b**). It is organized into characteristic densely packed DNA foci – the chromocenters – that are enriched in trimethylated lysine 9 of histone H3 (H3K9me3) and the H3K9me3-binding heterochromatin protein 1 (HP1) as seen by DAPI and immunofluorescence (IF) staining (**Fig. 1b**). PCH domains of several chromosomes coalesce within a chromocenter, whereas the neighboring minor satellites are localized adjacently as separate entities (Guenatri *et al.* 2004). Notably, the chromocenter compaction observed by DAPI staining persists upon loss of the heterochromatin H3K9 methyltransferase SUV39H and concomitant loss of H3K9me3/HP1 (Peters *et al.* 2001).

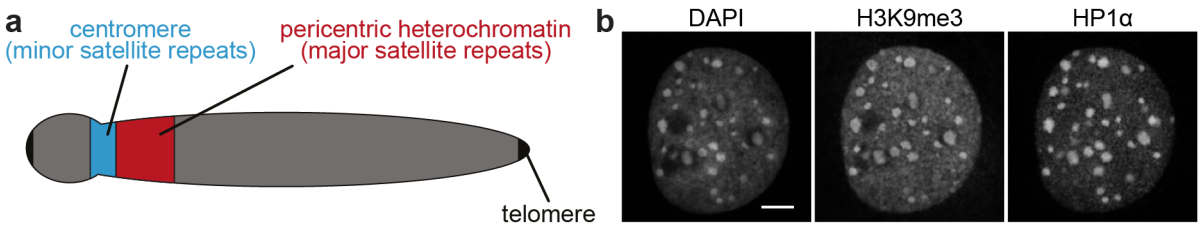

Figure 1: Pericentromeric major satellite repeats are organized into distinct nuclear subcompartments in mouse embryonic fibroblast cell lines. (**a**) Schematic representation of the linear arrangement of centromere and pericentromere in a condensed mitotic acrocentric mouse chromosome. (**b**) Confocal laser scanning microscopy (CLSM) images of interphase nuclei where PCH forms distinct, densely packed nuclear subcompartments that are trimethylated at H3K9 and enriched for HP1 (as shown by IF staining). Scale bars: 5 µm.

The centromere and the adjacent PCH play distinct roles in the organization of chromosomes within the nucleus throughout the cell cycle. The centromeres are involved in kinetochore formation during mitosis, whereas PCH is required for sister chromatid cohesion (see (Schalch & Steiner 2017) for a review). Consequently, disruption of PCH leads to mitotic defects in the form of chromosome missegregation (Guenatri *et al.* 2004). Taken together, intact PCH plays a crucial role in protecting the cells from genomic instabilities, which are associated with impaired viability and an increased tumor risk (Peters *et al.* 2001).

I.1.1 Suppressor of variegation histone methyltransferases are the central PCH components

H3K9me3 at PCH depends on the histone lysine methyl transferases (KMTs) SUV39H1 and SUV39H2. Their concurrent loss leads to upregulated transcription of the pericentromeric tandem satellite repeats, genomic instability and mitotic defects (Lehnertz *et al.* 2003, Martens *et al.* 2005, Peters *et al.* 2001). SUV39H enzymes constitute one of three families of eukaryotic methyltransferases that modify H3K9 (**Fig. 2a**). The other two families G9A and SET domain bifurcated (SETDB) have been described to prefer eu-

chromatic H3K9 as a substrate (Dodge *et al.* 2004, Schultz *et al.* 2002, Tachibana *et al.* 2002). G9A has, however, also been found to be important for heterochromatin silencing in mouse embryonic stem cells (mESCs) albeit through a mechanism that is independent of its KMT activity (Dong *et al.* 2008). In addition, the different KMTs preferentially catalyze different degrees of methylation – i. e. mono-, di- or trimethylation – at H3K9 which have different biological functions (Martens *et al.* 2005). SUV39H1 sets H3K9me3 while G9A mostly sets H3K9me1/2 (Peters *et al.* 2003). SETDB1 can catalyze all three methylation states, but is more efficient in setting H3K9me1/2 (Basavapathruni *et al.* 2016). In addition, its specificity and efficiency can be influenced by interaction with other proteins (Wang *et al.* 2003). The finding that SUV39H1 prefers mono- or dimethylated H3K9 as a substrate implies that it must cooperate with a mono- or di-KMT. Indeed, SUV39H1, G9A, GLP and SETDB1 participate in a multimeric complex that is also found at PCH (Fritsch *et al.* 2010). In addition, H3K9 monomethylation of cytoplasmic (i. e. non-nucleosomal) histones is set by PRMD3 and 16. Concurrent knock down of those two enzymes leads to a reduction of cytoplasmic H3K9me1-modified histones and disintegration of PCH (Pinheiro *et al.* 2012). SU(VAR)3-9 in *Drosophila melanogaster* was the first protein of the SU(VAR)3-9 family to be identified and characterized. X-ray-induced rearrangements cause a variety of phenotypes in *D. melanogaster* including the white eye variegation phenotype (Muller 1930). In this case, flies had patches of normal red pigmentation and patches of white indicating that the *white* gene was impaired, but not completely lost. In fact, the phenotype was caused by a rearrangement that placed the *white* gene in proximity of PCH and thereby led to its silencing. This effect is termed position effect variegation (PEV) and flies with a PEV phenotype were used to screen for enhancers or suppressors of variegation. SU(VAR)3-9 was identified as an antipodal dose-dependent **su**ppressor of **var**iegation (see (Elgin & Reuter 2013) for a review). Human and mouse SUV39H1 proteins display 95% amino acid sequence identity, but differ from fly SU(VAR)3-9 (42% and 41% identity for human and mouse, respectively) and yeast CLR4 (38% and 35% identity, respectively). However, there are three regions of sequence identity across all three species (**Fig. 2b**): the catalytical SET domain and its adjacent cysteine-rich regions as well as the chromo domain (CD) (Aagaard *et al.* 1999, O'Carroll *et al.* 2000). The SET domain is a highly conserved structure that is found in dozens of proteins in different organisms. Contrarily, the SET-domain flanking cysteine-rich regions are only found in a subgroup of KMTs including SUV39H1, SUV39H2, G9A and SETDB1 and might confer specificity for H3K9 (Kouzarides 2002). Extensive domain deletion and mutation studies of SUV39H1 revealed that the cysteine-rich regions or specific mutations within the SET domain abolished KMT activity (Rea *et al.* 2000). In addition, the very N-terminus of the protein (amino acids 1-44) and the adjacent CD are required for heterochromatin targeting (Melcher *et al.* 2000). The specific binding of the CD to H3K9me2/3 was shown *in vitro* (Wang *et al.* 2012) and probably contributes to SUV39H1

targeting *in vivo*. Lastly, various PTMs of SUV39H1 have also been implicated in its activity and cell cycle-dependent regulation. Phosphorylation of SUV39H1 within its SET domain reduces its gene silencing potential, is set during G1/S transition and persists until the end of mitosis (Firestein *et al.* 2000, Park *et al.* 2014). Upregulation of satellite transcripts is also observed upon methylation of SUV39H1 by the SET domain-containing protein SET7/9 (Wang *et al.* 2013). In contrast, deacetylation of lysine 266 within the SUV39H1 SET domain by SIRT1 has the inverse effect as it enhances its KMT activity (Vaquero *et al.* 2007). Last, SUV39H1 has also been found to be sumoylated between aa114-140 which might have implications for its interaction with the SUMO ligase UBC9 (Maison *et al.* 2016).

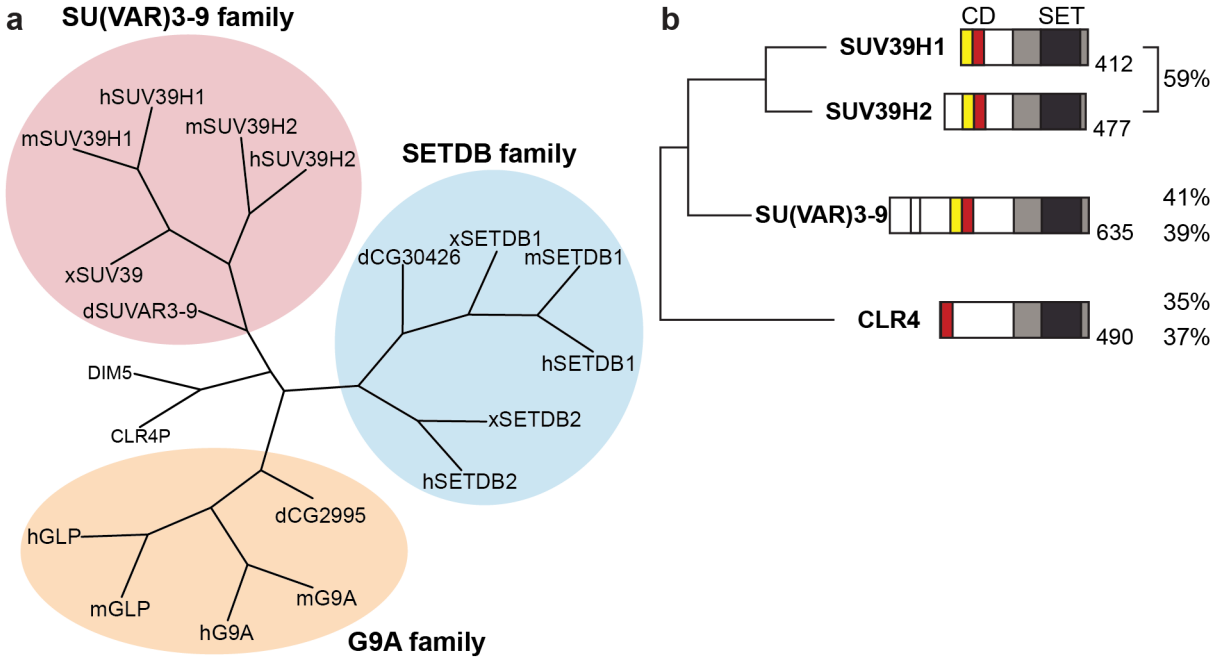

Figure 2: SUV39H and its relation to other KMTs. (**a**) Phylogenetic tree of the three families of eukaryotic KMTs. Homologues from *Homo sapiens* (h), *Mus musculus* (m) *Xenopus laevis* (x) and *Drosophila melanogaster* (d) are shown. DIM5 and CLR4P are KMTs from *Neurospora crassa* and *Schizosaccharomyces pombe*, respectively. Adapted and modified from (Huisinga *et al.* 2006). (**b**) Domain structure of fly SU(VAR)3-9 and its mammalian and yeast homologues. High sequence similarity is observed in the CD (red), the catalytic SET domain (black) and its flanking cysteine-rich regions (grey). The mammalian N-terminal domain (yellow) is also found in SU(VAR)3-9. The displayed mammalian SUV39H proteins are those from mouse. Adapted and modified from (Aagaard *et al.* 1999, O'Carroll *et al.* 2000).

Even though only loss of both SUV39H isoforms leads to genomic instability and susceptibility to tumorigenesis (Peters *et al.* 2001), SUV39H2 has been studied much less. In mice, SUV39H1 and SUV39H2 show 59% amino acid sequence identity (**Fig. 2b**). The most remarkable difference between the two proteins is the 82 aa N-terminal extension of SUV39H2 that contains ~20% basic amino acids (O'Carroll *et al.* 2000). Recent studies showed that the SUV39H2 substrate specificity differs from that of SUV39H1 (Ku-

dithipudi *et al.* 2017, Schuhmacher *et al.* 2015). Functional implications of this difference are largely unknown and might involve differential methylation of non-histone targets (Kudithipudi *et al.* 2017). The basic extension of SUV39H2 has also been shown to bind RNA (Velazquez Camacho *et al.* 2017), which might contribute to its increased immobilization both at PCH (∼90% vs. ∼20% for SUV39H1) and euchromatin (∼76% vs ∼3% for SUV39H1) (Muller-Ott *et al.* 2014). Taken together, SUV39H2 appears to be tethered more tightly to chromatin possibly also by other factors compared to SUV39H1, which could indicate that the functions of the SUV39H isoforms are not completely redundant and that they are also relevant outside of PCH.

I.1.2 SUV39H interacts with several factors to ensure PCH stability

SUV39H is a central component of a larger protein complex that forms specifically at PCH. Apart from its possible direct chromatin interaction via H3K9me2/3 mentioned above, SUV39H1 has been shown to interact with different isoforms of HP1 both *in vitro* (Maison *et al.* 2016, Melcher *et al.* 2000) and *in vivo* (Krouwels *et al.* 2005, Muller-Ott *et al.* 2014). As SUV39H enzymes set H3K9me3, they establish more binding sites for HP1, which contains a CD that binds H3K9me3 (Bannister *et al.* 2001). It has thus been proposed that SUV39H, HP1 and H3K9me3 are the core components of a network that is able to propagate the heterochromatic state linearly along the DNA fiber (Nakayama *et al.* 2001). Linear spreading would result in expanding the silenced chromatin state across the entire genome unless heterochromatin boundaries such as DNA binding sites for specific proteins, incorporated histone variants or antagonistically modified nucleosomes exist (see (Wang *et al.* 2014) for a review on limiting heterochromatin spreading). In an alternative model, heterochromatin spreading is assumed to initiate from sparsely distributed initiation complexes. Those are characterized by tight SUV39H-binding to PCH chromatin. From those complexes, H3K9me3 spreads in 3D and is inherently confined by the constant activity of counteracting demethylases (Erdel & Greene 2016, Muller-Ott *et al.* 2014). This model is supported by theoretical studies implicating that nearest neighbor spreading is not sufficient to establish stable domains (Dodd *et al.* 2007) and that a combination of *cis* and *trans* regulation are relevant for the regulation and memory of active or repressed states (Berry *et al.* 2017). Apart from the core PCH components, other factors are specifically enriched or active at PCH (**Fig. 3**). Downstream of SUV39H are the SUV4-20H KMTs that set trimethylation at lysine 20 of histone H4 (H4K20me3). SUV4-20H activity is directed towards PCH via H3K9me3/HP1 (Schotta *et al.* 2004). However, if H4K20me3 is functionally relevant for PCH formation and/or maintenance is under debate (Sakaguchi *et al.* 2008). SUV39H1, but not SUV39H2, has also been found to promote HP1α sumoylation through interaction with the SUMO ligase UBC9, thereby providing an additional targeting mechanism of HP1α to PCH (Maison *et al.* 2016). Last, DNA methylation plays an important role at PCH. In particular, seven of

the eight cytosine-phosphate-guanine (CpG) sites (i. e. ∼87%) found within the consensus sequences of major satellite repeats are frequently methylated (Horz & Altenburger 1981, Manuelidis 1981). The 5-methyl-cytosine (5mC)-binding domain (MBD) proteins MBD1 and MECP2 are strongly enriched at PCH, too (Muller-Ott *et al.* 2014). SUV39H1 and HP1β also interact with the DNA methyltransferases DNMT3A and DNMT1 establishing a direct link between the H3K9 and DNA methylation pathway (Fuks *et al.* 2003a).

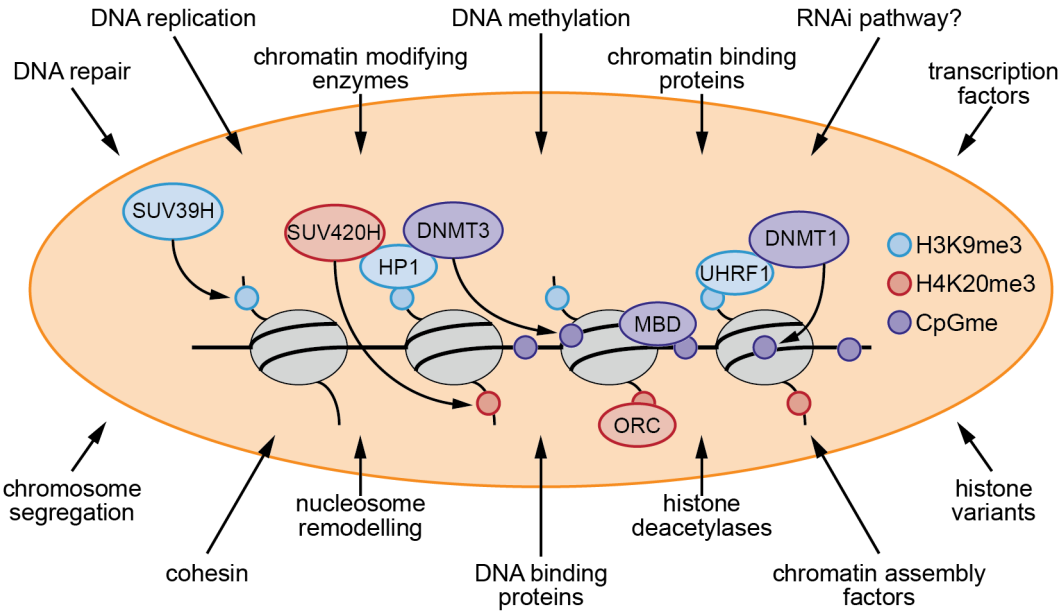

Figure 3: Overview of mammalian heterochromatin factors. Histone modifications – namely H3K9me3 and downstream H4K20me3 – cooperate with the DNA methylation pathway at constitutive heterochromatin such as PCH in mammals. Stabilized heterochromatin usually involves the concerted activity of 'reader' proteins that recognize modified histone tails or DNA (e. g. HP1 or MBDs) and 'writer' enzymes that set the respective modifications (e. g. SUV39H or DNMTs). 'Eraser' enzymes (not depicted here) remove modifications and are implicated in the regulation of domain sizes (Erdel & Greene 2016). Various interactions between involved proteins have been observed (see text for details). UHRF1: ubiquitin-like PHD and RING finger domain-containing protein 1 (a ubiquitin-protein ligase). ORC: origin of replication complex. CpGme: methylated CpG. Figure adapted and modified from (Saksouk *et al.* 2015).

The interdependencies of the SUV39H/H3K9me3/HP1 histone modification complex and the DNMT/5mC/MBD DNA methylation complex appear to differ between cell types. In mESCs, the histone modification complex directs DNA methylation (Lehnertz *et al.* 2003) and cells can proliferate and retain their stem cell characteristics in the absence of DNMTs and CpG methylation (Tsumura *et al.* 2006). However, quantification of histone marks at PCH using mass spectrometry showed that H3K9me3 and H4K20me3 are reduced upon loss of all DNMTs (Saksouk *et al.* 2014). In addition, Saksouk and colleagues showed that the facultative heterochromatin pathway centered around H3K27me3 is increased when DNA methylation is lost indicating that compensatory effects might be active (Saksouk *et al.* 2014). H3K9me3 seems thus to be upstream

of DNA methylation in mESCs, although the latter can modulate the former and plays a role in suppressing the alternative H3K27me3-centered heterochromatin route. In contrast, DNA methylation levels were found to be unaffected by loss of SUV39H in MEFs (Arand *et al.* 2012). DNA demethylation by 5-aza-2'deoxycytidine (5-aza-C) treatment only slightly increased SUV39H1 mobility (Krouwels *et al.* 2005) indicating that the PCH domain remains mostly intact even in the absence of DNA methylation. Thus, whether or not the histone methylation pathway at PCH is dependent on the DNA methylation pathway remains questionable.

I.2 Histone post-translational modifications and their relation to transcription

Post-translational modifications of histone tails have long been accepted to be critical for a variety of cellular functions including mitosis and meiosis, histone deposition and gene regulation (Strahl & Allis 2000). The repressive histone methylations at lysine 9 and lysine 27 of histone 3 demarcate constitutive and facultative heterochromatin, respectively. In contrast, histone tails at actively transcribed euchromatin are generally acetylated at lysines 9 and 14 (H3K9/14ac) among others. However, lysine methylation, e. g. at lysine 4 of histone 3, can also be attributed to transcriptionally active chromatin. Altogether, histone acetylation and methylation as well as even more modifications such as phosphorylation and ubiquitinylation constitute an additional layer of specific binding sites for effector proteins apart from the underlying DNA sequence (see (Lawrence *et al.* 2016) for a review). The complexity of possible combinations of multiple marks can be reduced to a limited number of chromatin states by computational methods (Baker 2011). Some of the modifications show strong localization at distinct genomic loci and have been attributed to distinct functions. Trimethylation of histone 3 lysine 4 (H3K4me3) and acetylation of H3K9 & H3K14 for instance are strongly enriched in the vicinity of the transcription start site (Liang *et al.* 2004) and play a fundamental role in transcription initiation (Guenther *et al.* 2007). In contrast, trimethylation of histone 3 lysine 36 (H3K36me3) is spread across the entire gene body (Barski *et al.* 2007) and implicated in elongation (Kizer *et al.* 2005). Other modifications such as H3K9me3 display a much less well-defined pattern across the genome (Bulut-Karslioglu *et al.* 2014, Molitor *et al.* 2017) and more diverse functional implications. Despite being generally attributed to silencing chromatin states (Bannister *et al.* 2001, Martens *et al.* 2005, Nielsen *et al.* 2001), H3K9me3 has also been described to play a role at actively transcribed genes as it associates with RNA polymerase II (Squazzo *et al.* 2006, Vakoc *et al.* 2005) and affects alternative splicing (Saint-Andre *et al.* 2011). The relationship between H3K9me3 and transcription thus remains elusive on the genome-wide scale.

I.2.1 SUV39H-dependent H3K9me3 outside PCH

Apart from its role at constitutive heterochromatin such as PCH, H3K9me3 has also been found to play a role in the regulation of gene expression. Several factors have been described to direct KMTs to the vicinity of promoters leading to enhanced H3K9me3 and repression of gene expression. Although H3K9me3-dependent gene regulation is mostly conferred by G9A, GLP and SETDB1 (Ayyanathan *et al.* 2003, Ogawa *et al.* 2002, Roopra *et al.* 2004, Schultz *et al.* 2002), SUV39H1 is directed to a subset of S-phase gene promoters by retinoblastoma (Rb) protein. The latter is in turn directed to the promoter set by binding to the E2F transcription factors, leading to H3K9me3 and HP1 binding (**Fig. 4a**). In this case, target genes – i. e. cyclin E and cyclin A – are upregulated upon loss of SUV39H, indicating that SUV39H KMTs can also regulate gene expression by directing promoter methylation (Nielsen *et al.* 2001, Vandel *et al.* 2001). Notably, SUV39H-dependent silencing of S-phase genes is particularly relevant in differentiating cells that permanently exit the cell cycle (Ait-Si-Ali *et al.* 2004). Other studies showed that unliganded thyroid hormone receptor also silences a reporter gene in a SUV39H-dependent manner (**Fig. 4a**) in *Xenopus oocytes* (Li *et al.* 2002) and that SUV39H1 and G9A are recruited by the EVI1 transcription factor to function as repressors of target transcription (Spensberger & Delwel 2008). In summary, promoter-directed SUV39H-dependent H3K9me3 can regulate the expression of target genes, but most studies have only considered a few targets and specific recruitment pathway, leaving the general contribution of SUV39H to gene expression unresolved.

Besides genes, SUV39H-dependent H3K9me3 also silences transcription of some repeat classes throughout the genome. Even though the cumulative enrichment of SUV39H-dependent H3K9me3 is rather low at long interspersed nucleotide elements (LINEs) and endogenous retroviruses (ERVs) (Martens *et al.* 2005), the subgroup of intact, full length LINEs and ERVs shows strong enrichment for SUV39H-dependent H3K9me3 in mESCs (**Fig. 4b**). However, loss of SUV39H only derepresses LINE transcription, whereas ERV silencing seems to rely mainly on SETDB1 (Bulut-Karslioglu *et al.* 2014). H3K9me3 of ERVs depends on the presence of intact long terminal repeats (LTRs) and in agreement with this finding, transcriptional latency of the LTR-containing human immunodeficiency virus (HIV) has also been attributed to SUV39H, H3K9me3 and HP1 (du Chene *et al.* 2007). In committed cells, a reduction of SUV39H-dependent H3K9me3 at LINEs and ERVs is accompanied by an increase in DNA methylation at those sites, indicating that retrotransposon silencing is ensured differently once cells differentiate. In contrast, SUV39H-dependent H3K9me3 domains of 1-2 Mbp that contain genes and coincide with lamina-associated domains (LADs) have been observed in iMEFs, but not in mESCs (Bulut-Karslioglu *et al.* 2014). Despite its general association with transcriptional silencing, H3K9me3 is enriched in the gene body of actively transcribed genes compared to their promoter regions. This enrichment depends on elongating DNA-dependent RNA

polymerase II (pol II) and is lost when transcription ceases (Vakoc *et al.* 2005). Squazzo and collegues also found H3K9me3 to colocalize with pol II at promoters (Squazzo *et al.* 2006). However, it is not clear which KMT sets this 'active' H3K9me3 mark.

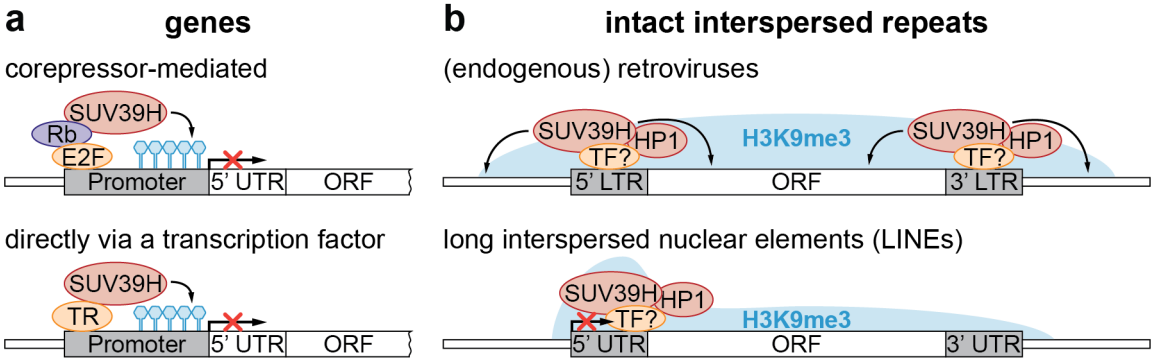

Figure 4: Transcriptional represssion through SUV39H-mediated H3K9me3 outside PCH. (**a**) Gene silencing via SUV39H-dependent promoter methylation (H3K9me3, blue). SUV39H is either recruited indirectly through a corepresser (e. g. Rb protein, top) or directly via a repressive transcription factor such as unliganded thyroid hormone receptor (TR, bottom). Based on results from (Li *et al.* 2002, Nielsen *et al.* 2001). (**b**) SUV39H-mediated H3K9me3 decorates intact interspersed repeats and represses transcription of full-length LINEs. How the repressor complex is targeted remains unclear, but might be realized through specific transcription factors (TFs). Based on results from (Bulut-Karslioglu *et al.* 2012).

I.2.2 Translating histone modifications into function

The information stored in modified histone tails generally requires so-called 'reader' domains for their translation into cellular functions. Methyl-binding CDs, acetyl-binding bromo domains and phospho-binding 14-3-3 proteins (**Fig. 5a**) are prototypical examples for a myriad of specific domains that have been described (see (Bannister & Kouzarides 2011, Musselman *et al.* 2012) for reviews). Notably, some of these domains are highly specific as they can distinguish between modifications at different residues of the histone tail (e. g. H3K9 vs. H3K4) and can also be sensitive to the degree of modification, e. g. in the case of methylation (Kim *et al.* 2006). Those domains can be part of proteins that serve an adaptor function by binding the modified histone tail and being a binding platform for other proteins (e. g. HP1). They can also be part of a catalytically active enzyme that has some chromatin function itself, e. g. SUV39H which serves a 'writer' function by setting a histone modification (**Fig. 5b**), thereby creating a positive feedback.

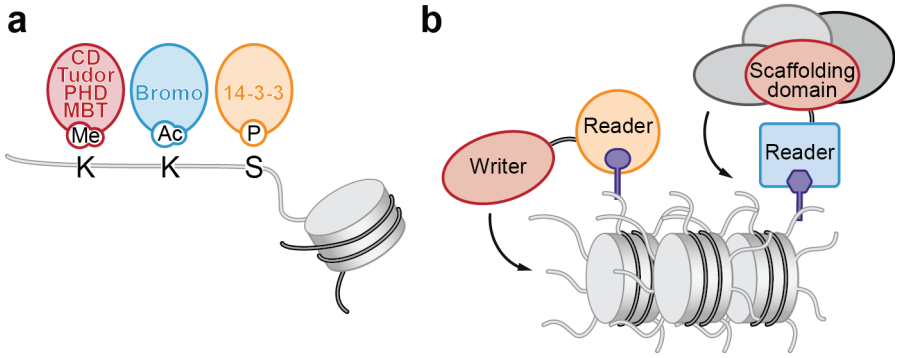

Figure 5: 'Readers' of histone modifications. (**a**) Modified histone tails are recognized by specialized reader domains. Chromo domains, Tudor domains, plant homeodomains (PHD), and malignant brain tumor (MBT) domains bind diverse methyllysines and Tudor domains can also bind methylarginines. Acetyllysines are bound by bromo domains (BD) whereas proteins from the 14-3-3 family bind phosphorylated serines or threonines (Musselman *et al.* 2012). Figure adapted from (Kouzarides 2007). (**b**) 'Reader' domains control downstream functions either by being part of catalytically active enzymes that might set histone modifications or by creating a binding platform for other effectors. Figure adapted from (Musselman *et al.* 2012).

Isolated reader domains for H3K9me3 have been used effectively to read out chromatin modifications in living cells (Muller-Ott *et al.* 2014) or even for chromatin immunoprecipitation (ChIP) instead of antibodies (Kungulovski *et al.* 2014, Su *et al.* 2014). In particular, the CDs of MPHOSPH8 and HP1, which share a conserved recognition mechanism (Li *et al.* 2011), and the ATRX-Dnmt3-Dnmt3L (hADD) domain of the human ATRX protein have been extensively tested on immobilized modified histone tails and in ChIP experiments and showed similar or even higher specificity compared to antibodies (Dhayalan *et al.* 2011, Kungulovski *et al.* 2014).

I.3 Gene regulation by controlling transcription

Transcription is a fundamental cellular process that is tightly regulated to preserve cell identity, prevent disease and appropriately react to environmental cues. It is one of the major control points to regulate gene expression, which can, in addition, be tuned at the level of RNA processing, transport, degradation and translation (Alberts *et al.* 1994, p. 403). In eukaryotes, there are three RNA polymerases that can be distinguished through their sensitivity to α-amanitin and have different functions. While RNA polymerase I transcribes large ribosomal RNAs and RNA polymerase III generates small RNAs such as transfer RNA, pol II is responsible for the transcription of protein-coding mRNA (Alberts *et al.* 1994, p. 367). Here, the focus is on pol II transcription, which generally proceeds in three main phases – initiation, elongation and termination – that are tightly regulated and control the overall RNA production rate.

I.3.1 Transcription by RNA polymerase II

Before transcription of a gene can start, the pol II complex consisting of twelve subunits associates with the general transcription factors on the promoter DNA to form the pre-initiation complex (PIC). In the stepwise assembly of the PIC (**Figure 6**), the general transcription factor TFIID binds the promoter DNA first. The canonical TFIID consists of a TATA box-binding protein (TBP) and several TBP-associated factors (TAFs) (Sainsbury *et al.* 2015). Only some human promoters contain the consensus TATA sequence near the transcription start site (TSS) (Yang *et al.* 2007), but several other core promoter elements that are bound by TAFs have also been identified (Baumann *et al.* 2010). TFIID binding leads to a bend in the DNA, which is stabilized by TFIIA and TFIIB binding the DNA and TBP. TFIIB then recruits pol II and the associated stabilizing factor TFIIF thereby completing the assembly of the core PIC. Pol II-binding of TFIIE and recruitment of TFIIH result in the closed PIC. A critical step for productive transcription is the opening of the promoter DNA (Wang *et al.* 1992), which depends on the ATPase activity of the TFIIH subunit XPB. In addition, TFIIH contains a kinase module including CDK7, which phosphorylates the C-terminal domain (CTD) of the largest subunit RPB1 of pol II and the helicase XPD, which is implicated in DNA repair (Sainsbury *et al.* 2015). Finally, the formation of the PIC is facilitated and stabilized by interactions with the multi-subunit mediator complex that interacts with several general transcription factors as well as DNA-binding transactivators and pol II (see (Conaway & Conaway 2011) for a review).

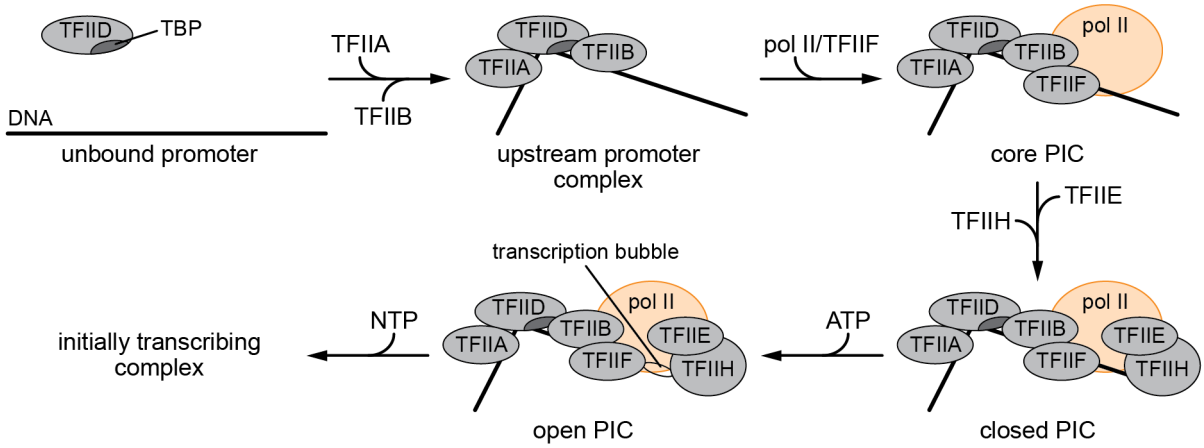

Figure 6: Formation of the transcription initiation complex at the core promoter. The general transcription factors TFIID, TFIIA, TFIIB, TFIIF, TFIIE and TFIIH associate with the core promoter DNA and pol II to form the pre initiation complex (PIC). DNA is melted under ATP consumption forming the transcription bubble that renders the single stranded DNA template accessible. Figure based on (Sainsbury *et al.* 2015).

Formation of the open PIC with the single stranded DNA template is not sufficient for pol II to enter its productively elongating state. In addition, pol II needs to detach

from the promoter-bound PIC components to move along the gene, i. e. achieve promoter escape, and overcome promoter-proximal pausing (Saunders *et al.* 2006). Upon open PIC formation, RNA synthesis is rapidly initiated in the presence of nucleotide triphosphates (NTPs). In this initially transcribing complex, short RNA products are synthesized and released, a process termed abortive initiation. This phenomenon is markedly reduced once the fourth nucleotide is added to the growing RNA chain when the transcription complex is committed to promoter escape (Holstege *et al.* 1997). After addition of the eighth nucleotide, the rate-limiting step of promoter escape – the transition to the early elongation complex – takes place. This coincides with the collapse of the upstream portion of the transcription bubble allowing the early elongation complex with the transcription bubble to move downstream along the DNA (Pal *et al.* 2005, Saunders *et al.* 2006). Another rate-limiting step is the pausing of pol II, which occurs shortly after promoter escape about 30-60 nucleotides downstream of the TSS (Kwak *et al.* 2013). Pol II pausing is controlled by negative elongation factor (NELF) and DRB-sensitivity-inducting factor (DSIF). It can also depend on core promoter features and associated sequence-specific transcription factors as well as nucleosomes (Jonkers & Lis 2015). To release pol II from the paused state, positive transcription elongation factor-b (P-TEFb) is required as its inhibition prevents pol II release and blocks transcription almost completely (Cheng *et al.* 2012, Henriques *et al.* 2013). P-TEFb consists of cycline T1 and CDK9 and phosphorylates NELF, DSIF and the CTD of pol II. CDK9-dependent phosphorylation of NELF and DSIF causes their dissociation or transformation into a positive elongation factor, respectively. Even though promoter-proximal pausing appears to be a regulatory event that represses transcription at first glance, it is in fact necessary for maintaining active transcription by competing with nucleosomes, which prevents the formation of repressive chromatin around the 5'-end of the gene (Gilchrist *et al.* 2010). Taken together, the level and duration of pausing and the pause release depend on the balance of pausing factors like NELF and DSIF and factors that recruit P-TEFb. This balance is strongly influenced by the activation of signaling pathways and could also be mediated by DNA looping that brings together promoters and enhancers thereby targeting additional cofactors possibly including enhancer RNAs to the pol II complex (reviewed in (Heinz *et al.* 2015)). Bromodomain-containing protein 4 (BRD4) is an exemplary activating factor that competes with an inhibitory complex for binding to P-TEFb and targets the latter to the promoter region by binding acetylated histones or the acetylated form of the transcription factor NF-κB among others. However, BRD4 is most likely not strictly required for pol II pause release, since the interplay of factors that mediate P-TEFb recruitment varies across different genes, cell types and stimuli (Jonkers & Lis 2015). After release from the promoter-proximal pause site, the RNA is productively elongated as transcription proceeds away from the 5'-end of the gene. Elongation rates differ across different genes (Fuchs *et al.* 2014) and are influenced by a variety of factors including histone PTMs, elongation factors, histone

chaperones, chromatin remodelers and DNA sequence composition (Jonkers & Lis 2015).

I.3.2 RNA polymerase II and histone post-translational modifications and the transcription of genes

The pol II CTD is conserved across different species and consists of a variable number of the heptapeptide YSPTSPS that is subject to PTM throughout the transcription process (reviewed in (Corden 1990)). The most prominent modifications are the phosphorylation of serine 5 and serine 2 (**Fig. 7a**, reviewed in (Egloff *et al.* 2012)). The modification status of the CTD influences its conformation and its interaction with enzymes involved in co-transcriptional nuclear processes such as mRNA biogenesis and chromatin modification (reviewed in (Meinhart *et al.* 2005) and (Buratowski 2009)). CTD serine 5 is phosphorylated early, namely in the PIC, by the TFIIH-subunit CDK7 and recruits the 5' mRNA capping machinery (Fabrega *et al.* 2003). It is also presumed to contribute to establishing H3K4me3 around the promoter region by recruiting MLL1 and/or SET1 complexes (Buratowski 2009). As pol II associates with P-TEFb in the proximity of the promoter to transition to productive elongation, CDK9 can phosphorylate CTD serine 2. While phosphorylation of serine 5 is reduced as elongation progresses, serine 2 phosphorylation is also set by CDK12 (Bartkowiak *et al.* 2010) and increases along the gene body (Heidemann *et al.* 2013). After transcription through the poly(A) site, pol II is paused, which allows for efficient transcription termination. Notably, transcription termination is also controlled by CTD modifications, in particular by phosphorylation of serine 2 and dephosphorylation of tyrosin 1 (see (Porrua & Libri 2015) for a review). Phosphorylation patterns of the CTD are controlled by the kinases mentioned above as well as phosphatases, such as SSU72, RPAP2, FCP1 and others. The latter dephosphorylates serine 2 at the end of transcription thereby recycling pol II for a new round of transcription. SSU72 and RPAP2 dephosphorylate serine 5 as the transcription machinery productively elongates the RNA resulting in the characteristic enrichment of serine 5 phosphorylation at the 5'-end of the gene (Egloff *et al.* 2012).

Specific histone modifications are typically found along actively transcribed genes (**Fig. 7b**) and influence pol II elongation rates. They also regulate co-transcriptional processes such as splicing or chromatin remodeling by constituting a binding platform for the respective complexes that frequently contain histone modification reader domains as described above. Histone acetylation directly influences pol II elongation as it disrupts the electrostatic interaction between negatively charged DNA and positively charged histones thereby facilitating the passage of pol II (Hong *et al.* 1993, Smolle & Workman 2013). Thus, histone acetylation is a typical pre-transcriptional mark (Venkatesh & Workman 2015) that has to be removed after passage of the transcription machinery to suppress aberrant transcription initiation (Carrozza *et al.* 2005). In contrast, methylation of H3K4 has a rather indirect effect by constituting a binding platform for other histone modifiers,

chromatin remodelers, DNA methyltransferases and others. H3K4me is set by SET1 and/or MLL KMTs that interact with the pol II CTD phosphorylated at serine 5 (Ng *et al.* 2003), but also depend on monoubiquitinated H2B (H2Bub), which is another typical histone mark associated with active transcription (Smolle & Workman 2013). H2Bub prevents the compaction of chromatin into higher order structures (Fierz *et al.* 2011) and promotes the reassembly of nucleosomes after the transcription machinery has passed through. It thereby favors a transcriptionally active state and intact chromatin after transcription. In addition, methylation of H3K36 is a typical mark that is associated with actively transcribed genes. While several H3K36me1/2 KMTs are found in higher eukaryotes, SETD2 is responsible for virtually all H3K36me3 (Edmunds *et al.* 2008). SETD2 interacts with pol II during elongation, possibly via the pol II CTD phosphorylated at serine 2 as has been shown for its yeast ortholog SET2 (Smolle & Workman 2013). Last, H3K9me3 is also found at actively transcribed genes, particularly on exons. Elevated H3K9me3 and HP1γ levels are associated with a decreased pol II elongation rate thereby favoring inclusion of alternative exons (Saint-Andre *et al.* 2011). Taken together, histone modifications and their 'readers' and 'writers' are important factors in the regulation of gene expression as well as co-transcriptional processes such as pre-mRNA splicing. They therefore contribute to the ordered disassembly and reassembly of chromatin in the context of transcription (see (Venkatesh & Workman 2015) for a review on transcription-associated histone exchange).

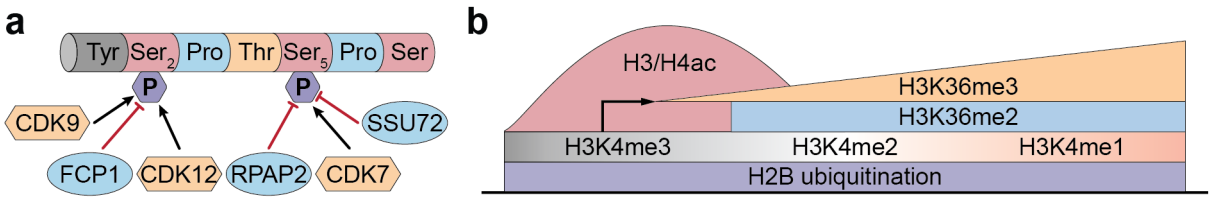

Figure 7: Post-translational modifications of the RNA polymerase II C-terminal domain and of histones across actively transcribed genes. (**a**) Phosphorylation of serine 2 and serine 5 are the most prominent modifications found on the pol II CTD. The most prominent enzymes that set (yellow) and remove (blue) those modifications are depicted, but more enzymes could play a role (e. g. CDK8 and CDK13 as well as the small CTD phosphatases). Figure adapted and modified from (Saunders *et al.* 2006) and also based on (Egloff *et al.* 2012). (**b**) Most prominent histone marks across a gene. 5'-ends predominantly harbor histone acetylation and H3K4me3 while H3K36me3 increases towards the 3'-end. Figure adapted and modified from (Saunders *et al.* 2006).

I.3.3 The viral transactivator VP16

The herpes simplex virus type 1 transcription factor VP16 has been used in numerous studies to artificially activate transcription and observe cellular processes associated with strong activation (e. g. (Tumbar *et al.* 1999)). In particular, its C-terminal transactivation domain (TAD) is a potent transcription activator that has been fused to a variety of proteins, e. g. DNA-binding proteins (e. g. (Janicki *et al.* 2004)), for a targeted activation of genes. The VP16 TAD binds to numerous general transcription factors, including some of the TAF-subunits of TFIID (Hirai *et al.* 2010). When bound to a TAF, the TAD assumes an α-helical conformation, whereas it is an unstructured random coil in its unbound form (Uesugi *et al.* 1997). VP16 binding to general transcription factors and mediator components as well as its recruitment of histone lysine acetyltransferases (KAT) and chromatin remodelers to promoters (Hirai *et al.* 2010) underlines its importance in stimulating transcription initiation. The lack of sequence conservation among strong activators and their low affinity contacts to components of the transcription machinery led to the proposal of an alternative model of activating initiation. It suggests the destabilization of promoter nucleosomes as a primary mechanism followed by low-affinity recruitment of the transcription machinery (Erkina & Erkine 2016). In contrast to some other transcription activators such as SP1, that mainly stimulates initiation, or HIV Tat, that mainly stimulates elongation, the VP16 TAD strongly enhances both initiation and elongation (Blau *et al.* 1996). Recruitment of VP16 TADs to repetitive heterochromatic arrays can be observed by live cell microscopy and leads to large-scale decondensation of the locus in a way that is independent of transcription (Carpenter *et al.* 2005, Tumbar *et al.* 1999). It should be noted, however, that the phenotype of those loci displays a high cell-to-cell variability as 32-39% of the cells retained compact structures comparable to the negative control (Tumbar *et al.* 1999). If and to what extent this decompaction depends on the ability of the VP16 TAD to recruit KATs such as GCN5, P300 or PCAF, that result in hyperacetylated histones at the locus, is unclear. Microscopy studies such as the one described above rely on bacterial DNA sequences at relatively high copy number for recruiting and visualizing the VP16 TAD and on reporter constructs such as the phage MS2 RNA that can be visualized by fluorescently labeled MS2 coat protein (MCP) (Shav-Tal *et al.* 2004). Those reporter cell lines are valuable systems as they allow for live cell experiments over time in single cells as well as extensive characterization of factors bound to the reporter locus in the inactive and active state (Janicki *et al.* 2004, Rafalska-Metcalf *et al.* 2010). In other studies, the very potent VP64 activation domain, which consists of four copies of the VP16 minimal activation domain (Beerli *et al.* 1998), was recruited to endogenous loci. This allowed to precisely trigger the expression of certain genes, but sacrificed single cell resolution (Gao *et al.* 2014, Konermann *et al.* 2015, Polstein *et al.* 2015). Altogether, VP16 is a strong transcriptional activator that has been widely used to activate both reporter and endogenous genes.

I.4 Targeting effector proteins

Directing effector proteins such as transcriptional activators, repressors or histone modifiers to certain genomic loci is a powerful tool to study the effector function in the native chromatin environment in living cells. Effector proteins can either be fused directly to DNA-binding domains (DBD) or to adaptor proteins that interact with their respective DBD-fused partner proteins under certain conditions.

I.4.1 DNA-binding domains and their use in living cells

One exemplary system to target DNA loci *in vivo* is the *lac* operator/repressor system from *Escherichia coli*, that has been adapted for use in eukaryotic cells (Fieck *et al.* 1992). In *E. coli*, the *lac* repressor LacI binds to DNA in the presence of the favorable glucose as energy source, thereby inhibiting transcription of genes involved in lactose metabolism. If glucose is lacking and lactose is used as an alternative energy source, LacI-*lac* operator (*lac*O) binding is abrogated and transcription activated. LacI binds to a single operator sequence as a homodimer, but can also form homotetramers and bind two operator sequences resulting in DNA looping (reviewed in (Swint-Kruse & Matthews 2009)). To monitor a certain part of a genome, multiple copies of the *lac*O sequence are integrated into the genome and visualized by transfecting fluorescently labeled LacI (Robinett *et al.* 1996). As described above, this system is frequently used to study large-scale changes in chromatin structure, e. g. decondensation upon VP16 recruitment.

Another bacterial DNA recruitment system is the *tet* operator/repressor system found in gram-negative bacteria that is responsible for bacterial resistance to tetracyclines (reviewed in (Ramos *et al.* 2005)). The *tet* repressor TetR strongly binds to the *tet* operator *tet*O as a dimer which inhibits the transcription of the resistance gene *tetA*. In the presence of tetracycline-magnesium complexes, TetR affinity for *tet*O is reduced by nine orders of magnitude, which triggers TetR dissociation and transcription of the resistance gene (Lederer *et al.* 1995, Orth *et al.* 2000). TetR has been used in a variety of studies in eukaryotic systems both as a direct repressor probably by sterically hindering pol II binding or as an activator when fused to an activation domain like the VP16 TAD (Berens & Hillen 2003). In the latter case, *tet*O repeats are placed next to a minimal promoter followed by a gene of interest such as luciferase, whose expression can be controlled over five orders of magnitude using the TetR/*tet*O system (Gossen & Bujard 1992). However, substantial drawbacks of this inducible targeting system exist. They include promoter- or integration site-dependent leakiness, constant exposure to tetracycline to ensure dissociation of the activator in the non-activated case and slow induction kinetics due to the required wash out of the inducer. All issues can be overcome or at least be reduced by using reverse TetR systems that only bind *tet*O in the presence of tetracycline or doxycycline. Still, inappropriate expression of the gene of interest in (i. e. "leakiness") has also been observed

when using reverse TetR. This unwanted activity can be attributed to basal activity of the minimal promoter or residual binding of the activator in the absence of doxycycline (Berens & Hillen 2003). Doxycycline-inducible systems can be used to induce activator recruitment within minutes (Normanno *et al.* 2015) and have been widely used in live cell studies dissecting the events centered around transcription activation by VP16 (Janicki *et al.* 2004, Rafalska-Metcalf *et al.* 2010).

Last, effectors can be targeted sequence-specifically to endogenous loci by using transcription activator-like effectors (TALEs) or the CRISPR/Cas9 system. TALEs (see (Boch & Bonas 2010, Zhang *et al.* 2014) for reviews) are based on a bacterial system that infects plants and activates the transcription machinery of the host resulting in the expression of genes involved in pathogenesis. TALEs can bind specific DNA sequences through their repeat domain that contains a hypervariable diresidue. The nature of this diresidue determines the DNA base that is recognized thereby providing a modular system with a direct protein-DNA binding code that can be tailored to specific target sequences (Boch *et al.* 2009). In contrast, the CRISPR/Cas9 system is a central component of the bacterial immune system against viruses (Horvath & Barrangou 2010) and relies on a guide RNA rather than a specific protein sequence for DNA sequence recognition (Cong *et al.* 2013, Mali *et al.* 2013). The guide RNA contains a specific CRISPR RNA sequence that is designed to be complementary to a 23 bp genomic target sequence allowing to target virtually any genomic site of the form $GN_{20}GG$ (Mali *et al.* 2013). The other component, Cas9, is a nuclease that cuts the genomic target sequence rendering the CRISPR/Cas9 system very efficient for homologous recombination-based genome editing. Using fusion proteins containing an endonuclease-deficient Cas9 version, allows this system to be used to target a fluorescent protein (CRISPR imaging, (Chen *et al.* 2013)) or effectors such as activators to specific genomic loci (Hsu *et al.* 2014, Konermann *et al.* 2015). Both TALE- and CRISPR/Cas9-targeted gene activation have been shown to be effective and highly specific rendering those tools valuable for functional studies of gene activity (Gao *et al.* 2014, Polstein *et al.* 2015).

I.4.2 Controlling effector localization with light

Apart from targeting proteins by fusing them directly to DBDs that can be chemically inducible (see above), effector localization can also be controlled with light. One way of achieving light sensitivity is the attachment of light-sensitive probes to biomolecules which leads to "caging", i. e. rendering the molecule biologically inert (see (Ellis-Davies 2007) for a review). Upon illumination, the probes are cleaved resulting in the biologically active form of the molecule. However, the design and synthesis of those probes is complex and labor-intensive restricting its use. In contrast, optogenetic tools – i. e. light-sensitive, genetically encoded proteins – occur naturally in plants and corals and are easy to use (see (Tischer & Weiner 2014) for a review). Three popular plant-based optogenetic

systems are cryptochromes, light-oxygen-voltage (LOV) domains and phytochromes. In the presence of red (phytochrome B, PHYB) or blue (cryptochrome 2, CRY2, and LOV domains) light, the light-sensitive proteins undergo conformational changes that have different consequences, namely hetero- or homodimerization, oligomerization or uncaging of a functional group.

PHYB associates with the chromophore PCB and the complex changes conformation upon exposure to red light, resulting in a form that can heterodimerize with its binding partner PIF. Illumination with far red light reverses PHYB to its inactive conformation resulting in dissociation from PIF (Ni *et al.* 1999). LOV domains rely on the chromophore flavin and their conformation change upon blue light illumination. They can be used to relieve the autoinhibition of a fused catalytic domain (Wu *et al.* 2009), to expose a binding site for homo- or heterodimerization (Motta-Mena *et al.* 2014, Strickland *et al.* 2012) or to expose a localization signal triggering relocalization of a fused protein (Niopek *et al.* 2014, 2016). CRY2 does not depend on an additional chromophore and changes its conformation upon illumination with blue light. It interacts with its binding partner cryptochrome-interacting basic helix-loop-helix 1 (CIB1) in this switched conformation (Liu *et al.* 2008). Kennedy and colleagues have adapted this system for use in mammalian cell lines and showed that only the PHR domain of CRY2 and the N-terminal domain of CIB1 (CIBN) are required for heterodimerization (**Fig. 8a**). In this pioneering study, CIBN was fused to a cytoplasma membrane-anchored GFP and PHR was fused to mCherry (**Fig. 8b**). Upon illumination, mCherry colocalized with GFP showing that fusion constructs of CIBN/PHR with other proteins such as fluorescent domains do not interfere with the conformation change (Kennedy *et al.* 2010), rendering the PHR/CIBN system suitable for live cell experiments. In addition, the interaction is reversible and CIBN and PHR dissociate within minutes after blue light illumination ceased.

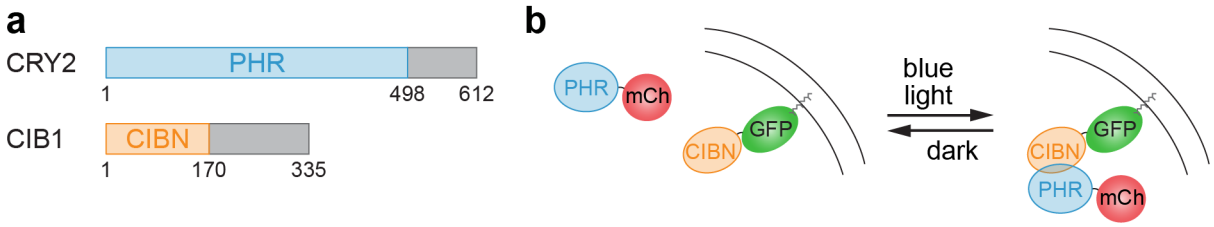

Figure 8: Components of an optogenetic system from *Arabidopsis thaliana*. (**a**) Schematic representation of CRY2 and CIB1 and the PHR and CIBN domains that are required for heterodimerization. The numbers correspond to the amino acid positions. (**b**) Principle of controlling the location of PHR-tagged proteins. The GFP is tethered to the cytoplasma membrane and mCherry colocalizes with GFP upon exposure to blue light. Within minutes in the dark, PHR-mCherry disperses again. mCh: mCherry. Figure adapted and modified from (Kennedy *et al.* 2010).

For the regulation of gene expression, inducible homo- or heterodimerization is particularly suitable to target effectors such as an activator or repressor of transcription to specific DNA loci (**Fig. 9a,b**). In the first case, the effector is fused to a LOV domain-containing protein that does not bind DNA in the dark. Upon illumination, a conformational change allows rapid homodimerization and DNA binding thereby inducing transcription (Motta-Mena *et al.* 2014). In the second case, a DBD is fused to one of the optogenetic interaction partners (e. g. CIBN) while the effector is fused to the other partner (e. g. PHR). Upon illumination, the effector colocalizes with the DBD analogous to the scheme depicted in **Figure 8b**. Notably, tagging the effector with either one of the optogenetic constructs has been shown to result in recruitment and activation of the respective gene (Konermann *et al.* 2013, Polstein & Gersbach 2015). Alternatively, an effector protein could be tagged with a nuclear localization signal (NLS) that is blocked by a LOV domain in the dark and exposed when illuminated resulting in translocation of the effector to the nucleus and subsequent DNA binding (Niopek *et al.* 2014) as depicted in **Figure 9c**.

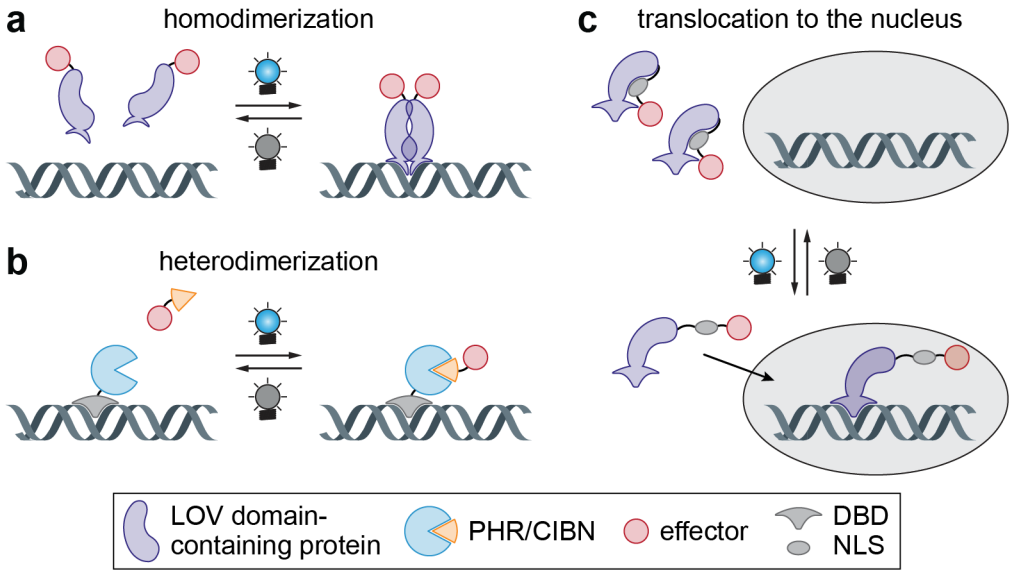

Figure 9: Strategies to control gene expression with light. Effectors such as activators or repressors of transcription can be recruited to DNA loci using different light-dependent strategies. (**a**) LOV domain conformational change can trigger homodimerization and DNA binding. (**b**) A DNA-bound optogenetic component can recruit its binding partner fused to an effector. (**c**) An NLS can be exposed through conformational change of a LOV domain leading to translocation to the nucleus and subsequent DNA binding. Figure adapted and modified from (Tischer & Weiner 2014).

Although a variety of optogenetic tools are available and have been used for transcription activation, transcription readouts can mostly not keep up with the rapid recruitment strategies since they rely on protein synthesis (e. g. luciferase) or cell lysis and PCR.

I.5 Scope of this thesis

Covalent modifications of histone tails are prominent marks that define chromatin states and control cell type specific gene expression programs and other genome activities. However, their establishment, maintenance and functional relevance in controlling cellular function are only partly understood. This thesis focuses on epigenetic silencing both at genes and other genomic elements such as pericentromeres and repeats. Furthermore, the kinetics of transcription activation of a locus that was initially in an inactive and heterochromatic state was dissected.

First, the question how *specificity* is conferred to silencing complexes and how silenced domains are *memorized* through cell division was addressed. Murine pericentric heterochromatin (PCH) was used as a model system for H3K9me3-dependent constitutive heterochromatin since it forms distinct nuclear patterns and can be distinguished from surrounding euchromatic regions by light microscopy. To that end, three potential mechanisms were explored: i) recognition of DNA sequence motifs in the major satellite sequence by specific transcription factors, ii) targeting of the SUV39H KMT through interaction with an upstream (epigenetic) factor and iii) 'bookmarking' of PCH sites via a factor that remains chromatin-bound throughout mitosis.

Second, the *redundancy* of epigenetic regulation was addressed in the PCH model system. It is often observed for numerous isoforms of histone modification 'writer' enzymes and probably confers stability to epigenetic states. However, the two isoforms of SUV39H show distinct protein mobility differences both at PCH and euchromatin (Muller-Ott *et al.* 2014) as well as different specificity profiles (Kudithipudi *et al.* 2017, Schuhmacher *et al.* 2015) suggesting that their functions might not be completely redundant. A systematic analysis of SUV39H1/2 single and double null as well as wild type immortalized mouse embryonic fibroblast (iMEF) cell lines was conducted to assess the redundancy of the two isoforms for maintaining the integrity of PCH. This analysis was extended to a comparative genome-wide ChIP- and RNA-seq study to gain a broader view on the relation of SUV39H-dependent H3K9me3 and transcription.

Third, the mechanism by which a heterochromatic locus can be *activated* was studied. What is the *cause or consequence* relation between histone modifications and transcription is one fundamental question that cannot be answered by steady-state analysis in functional knock-out cell lines. Accordingly, a time-resolved analysis unraveling the sequence of events involved in transcription activation was used to answer this 'chicken-or-egg' question. It was implemented by developing a light-inducible chromatin recruitment tool called BLInCR for **B**lue **L**ight-**In**duced **C**hromatin **R**ecruitment. A thorough characterization of its recruitment and reversibility properties was conducted for different effector proteins rendering BLInCR a versatile tool for time-resolved live cell microscopy studies. Finally, transcription was activated in the human U2OS 2-6-3 cell line with BLInCR to reverse the initially heterochromatic state of an array of reporter gene repeats. Fluores-

cently tagged live cell RNA and histone modification readers combined with rapid and reversible induction of transcription shed light on the kinetics of gene regulation and the stability of the heterochromatic state.

In the first two parts of the experimental analysis, the specificity and memory as well as the consequences of the loss of heterochromatin-associated H3K9me3 modification was dissected at endogenous loci. This was combined with a transcription activation analysis of a repressed reporter locus yielding mechanistic insights on silencing and activating heterochromatin. Heterochromatin is a barrier towards reprogramming of cell identity (Becker *et al.* 2016) and its deregulation is associated with disease (Benard *et al.* 2014). Hence, understanding the establishment and maintenance of this silenced chromatin state as well as processes that lead to derepression of transcription is fundamental.

II Results

II.1 Specificity and memory of endogenous silencing complexes

To elucidate how specificity is conferred to protein complexes that localize to PCH and how the heterochromatic state could be restored after it has suffered from replicative dilution during cell division, three possible targeting mechanisms were explored: i) via a sequence-specific factor, ii) via an (epigenetic) signal upstream of H3K9me3/SUV39H and iii) via a DNA sequence-independent 'bookmarking' factor that remains chromatin-bound throughout the cell cycle and thereby 'memorizes' the positions where PCH needs to be reestablished.

II.1.1 Targeting PCH through PAX transcription factors

PAX transcription factors have been proposed to specifically recognize PCH and to recruit other factors that are required for heterochromatin function (Bulut-Karslioglu *et al.* 2012). Fluorescently tagged PAX constructs previously described elsewhere (Bulut-Karslioglu *et al.* 2012) were transfected into NIH3T3 cells and their enrichment at PCH was analyzed in confocal laser scanning microscopy (CLSM) images (**Fig. 10a**). PAX7, which has a domain structure similar to that of PAX3, but is not usually expressed in MEFs (Bulut-Karslioglu *et al.* 2012), showed PCH enrichment in 33% of the cells. When transfecting PAX3-GFP, PCH enrichment was observed in only one cell (5%). Furthermore, neither PAX9-GFP nor PAX5-GFP was enriched at PCH (**Fig. 10a,b**). The latter has no consensus binding site within major satellite repeats and thus served as a negative control (Bulut-Karslioglu *et al.* 2012). The mobility of PAX3-GFP at PCH foci and in euchromatic regions was measured by fluorescence recovery after photobleaching (FRAP). To distinguish PCH and euchromatin in living cells, an NIH3T3 cell line stably expressing H2A-mRFP was used (Muller *et al.* 2009). The PAX3-GFP mobility was not significantly different in PCH compared to euchromatin. This was observed for cells with and without PAX3-GFP enrichment in heterochromatin (**Fig. 10c**). In general, PAX3 was very mobile and FRAP curves mostly fitted an effective diffusion model (**Table 1**), indicating that PAX3 is not a stably bound component of PCH initiation complexes (Muller-Ott *et al.* 2014).

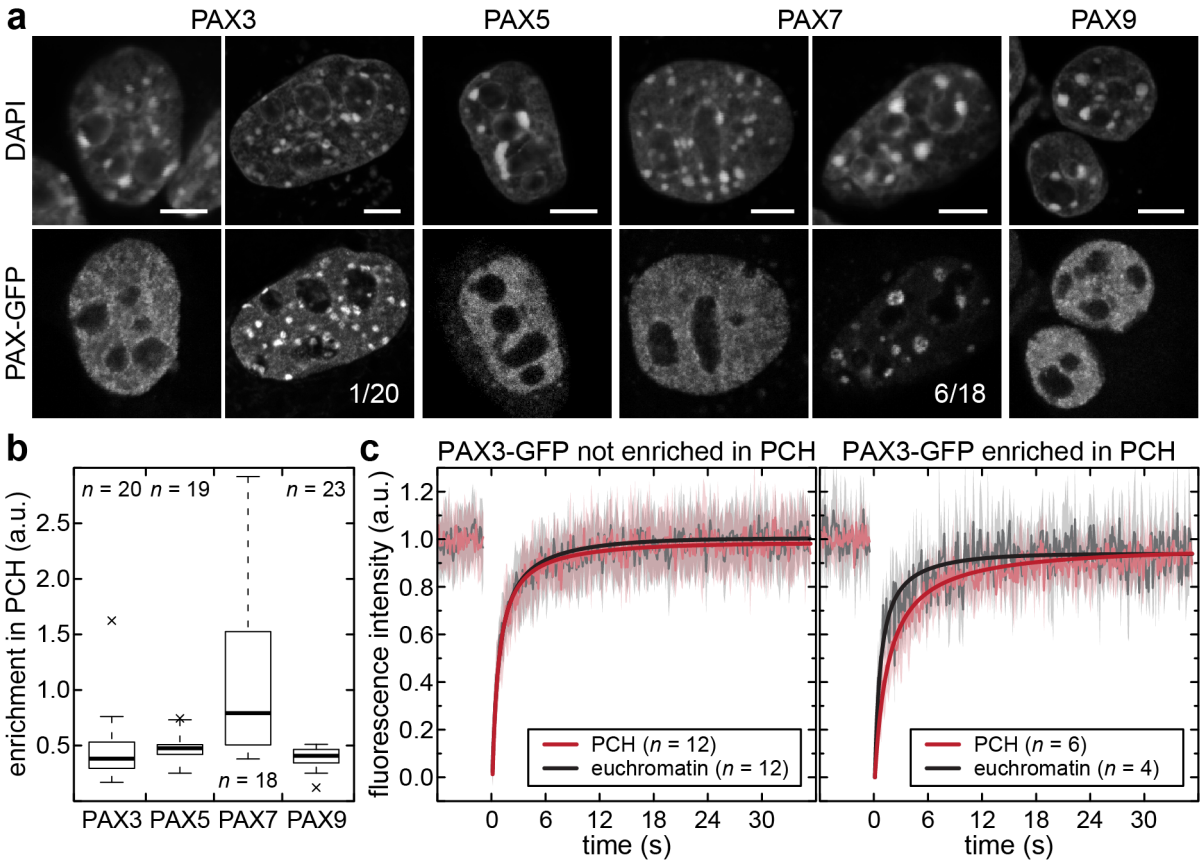

Figure 10: Enrichment and mobility of PAX transcription factors at PCH. (**a**) CLSM images of NIH3T3 cells transfected with PAX-GFP constructs, fixed and counterstained with DAPI (top). PAX7-GFP colocalized with DAPI-dense foci in 33% of the cells, PAX3-GFP only in 5%. PAX5- and PAX9-GFP enrichment at PCH was not observed. Scale bars: 10 µm. (**b**) Quantification of PAX-GFP enrichments at PCH. PCH/euchromatin ratios were normalized to chromatin density. (**c**) Fluorescence intensity traces after photobleaching at PCH foci and in euchromatic areas in cells without (left) or with (right) PAX3-GFP enrichment at PCH. Gray/pink lines represent averages of the experimental data, shaded areas are the respective standard deviations. Diffusion model fits are displayed as black/red solid lines. The respective fit parameters are listed in **Table 1**. PAX3- and PAX7-GFP images in (**a**) as well as the FRAP curves in cells without PCH enrichment in (**c**) were adapted from (Muller-Ott *et al.* 2014).

Table 1: PAX3-GFP mobility fit parameters.

	PCH			euchromatin		
	all	group 1	group 2	all	group 1	group 2
Pax3 enrichment in PCH	+/-	-	+	+/-	-	+
n	22	12	6	18	12	4
best fit model	d (19/22)	d (12/12)	d/rd (3/6)	d (17/18)	d (12/12)	d (4/4)
immobile fraction (%)	4.9±5.8	1.9±2.0	5.0±5.9	1.0±1.7	1.4±1.9	0.0±0.0
diffusion coefficient D (μm^2 s^{-1})	0.44±0.27	0.63±0.15	0.56±0.46	0.68±0.33	0.59±0.24	1.13±0.32

Single FRAP curves were fitted using FREDIS (Muller *et al.* 2009). Average curves of group 1 and 2 are depicted in **Figure 10c** (a fit of the average curves is displayed; fit parameters of the average were comparable to the single curve fits and are not listed here). Only cells that could be confidently assigned to the groups were used for the group analysis. The best fit models were either an effective diffusion model (d) or a model also accounting for binding of the factor to chromatin (reaction-diffusion model, rd). Mean and standard deviation of the single curve fits are listed for the immobile fraction and the diffusion coefficient.

II.1.2 MECP2 mobility was not influenced by SUV39H enzymes

The mobility of the methyl-DNA-binding protein and PCH component MECP2 was measured by FRAP in the presence and absence of SUV39H in wild type (wt) iMEFs and SUV39H double null (dn) iMEFs, respectively. GFP-MECP2 was transfected and FRAP was performed at PCH foci and euchromatic regions. PCH foci were clearly identified by enrichment of MECP2 in both wild type and SUV39H dn iMEFs (**Fig. 11a**). A fraction of GFP-MECP2 (∼15% in both wt and SUV39H dn iMEFs) was immobilized at PCH foci (**Fig. 11b**). This fraction was reduced to ∼3% in euchromatic regions in both cell types (**Table 2**). These results indicate that the mobility of GFP-MECP2 is different in PCH compared to euchromatin, but this difference did not depend on the absence or presence of SUV39H. GFP-MECP2 might thus act completely independently and upstream of SUV39H enzymes and could play a role in targeting the latter to PCH.

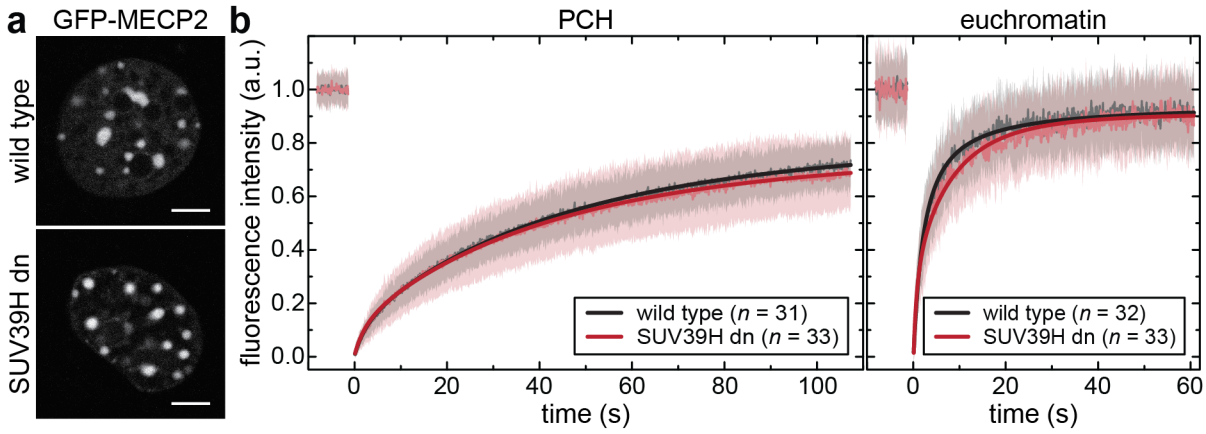

Figure 11: GFP-MECP2 mobility remained unchanged in the presence and absence of SUV39H enzymes. (**a**) CLSM images of wild type (top) and SUV39H dn (bottom) iMEFs transfected with GFP-MECP2, which was then enriched at PCH foci in both cell types. Scale bars: 5 µm. (**b**) Fluorescence recovery curves after photobleaching of PCH foci (left) or euchromatic regions (right) in wt or SUV39H dn iMEFs. Gray/pink lines represent averages of the experimental data, shaded areas are the respective standard deviations. Effective diffusion (euchromatin, wt) or reaction-diffusion model fits (all other curves) of the average curves are displayed as black/red solid lines. The respective fit parameters are listed in **Table 2**.

Table 2: GFP-MeCP2 mobility fit parameters.

	PCH		**euchromatin**	
	wild type	SUV39H dn	wild type	SUV39H dn
n	31	33	32	33
best fit model	rd (27/31)	rd (22/33)	d/rd (18 and 14/32)	rd/d (17 and 16/33)
immobile fraction (%)	13.9±7.7	16.6±10.0	3.1±3.7	2.6±4.4
diffusion coefficient D (μm^2 s^{-1})	0.013±0.004	0.012±0.004	0.15±0.1	0.16±0.08

Single FRAP curves were fitted using FREDIS (Muller *et al.* 2009). The best fit models were either an effective diffusion model (d) or a model also accounting for binding of the factor to chromatin (reaction-diffusion model, rd). Average curves for GFP-MECP2 mobility in wild type and SUV39H double null iMEFs at PCH and in euchromatic regions are depicted in **Figure 11b** (a fit of the average curves is displayed; fit parameters of the average were comparable to the single curve fits and are not listed here). Mean and standard deviation of the single curve fits are listed for the immobile fraction and the diffusion coefficient.

II.1.3 SUV39H1 colocalized with DNA throughout the cell cycle

To assess if SUV39H1 dissociates from chromatin during cell division, living NIH3T3 cells stably expressing SUV39H1-GFP upon induction with doxycycline were imaged. Hoechst 33342 was used to counterstain DNA and evaluate the cell cycle phase. As shown in **Figure 12**, SUV39H1-GFP (top) colocalized with Hoechst 33342 (middle) during interphase and early prophase as well as during anaphase, telophase and cytokinesis. From prophase through metaphase until early anaphase, a substantial amount of SUV39H1-GFP was evenly distributed throughout the nucleus or cytoplasm. However, a fraction of SUV39H1-GFP was enriched at sites of condensed DNA throughout the cell cycle, possibly constituting a 'bookmarking' factor for the reestablishment of PCH domains after cell division.

In summary, sequence-specific PAX transcription factors are unlikely to contribute to targeting of silencing complexes. Rather, MECP2 and DNA methylation could be upstream factors that direct SUV39H enzymes to mouse PCH. Alternatively or in addition, a fraction of the silencing machinery could remain chromatin-bound thereby contributing to both specificity and memory of PCH.

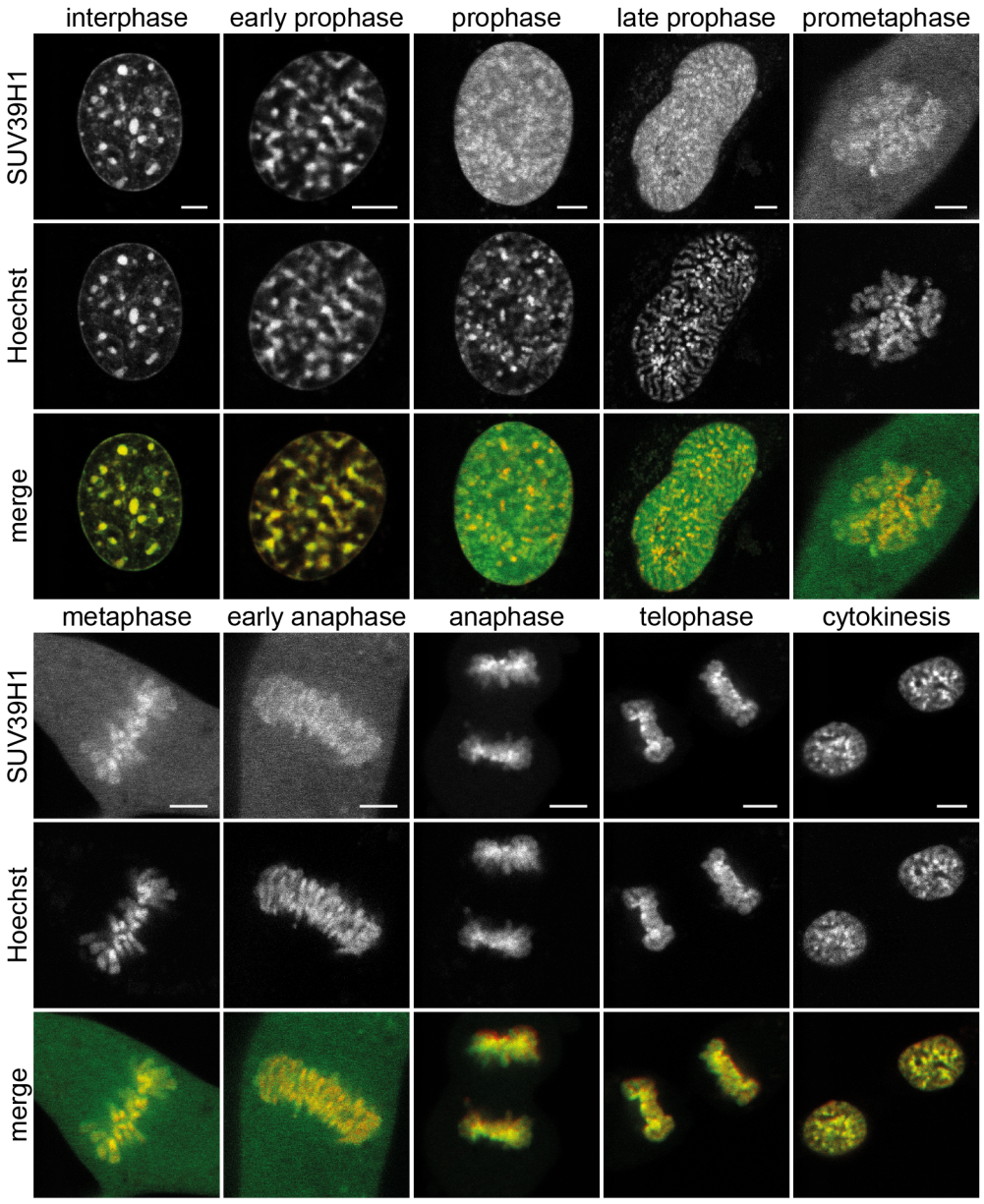

Figure 12: SUV39H1-GFP colocalized with DNA throughout the cell cycle. NIH3T3 stably expressing SUV39H1-GFP were incubated with the DNA stain Hoechst 33342 prior to CLSM imaging. Depicted are different cells from different cell cycle phases. A substantial amount – but not all – of SUV39H1-GFP dissociated from DNA-dense regions during prophase, metaphase and early anaphase. Hoechst is represented in red in the merge, so that colocalization with green SUV39H1-GFP appears yellow. Scale bars: 5 µm. Images adapted from (Muller-Ott *et al.* 2014).

II.2 The apparent redundancy of SUV39H histone methyltransferases in transcriptional silencing

Both SUV39H isoforms are significantly immobilized at PCH, but the fraction of immobilized protein is different at PCH and euchromatin (Muller-Ott *et al.* 2014) suggesting that SUV39H1 and SUV39H2 might have different functions. To dissect the role of the SUV39H isoforms, iMEFs with functional knock-out of either one or both of the isoforms were analyzed with respect to their H3K9me3 levels as well as their transcription activity.

II.2.1 Characterization of iMEF cell lines

The SUV39H1 or SUV39H2 single null (sn) cell lines were provided by Gunnar Schotta (LMU Munich) and were generated by a *lacZ* knock-in using homologous recombination as described in (Peters *et al.* 2001). An overview of the cell lines used for the respective experiments is given in **Table 3**.

Table 3: Overview of iMEF cell lines used in the experiments described below.

name	SUV39H1	SUV39H2	microscopy	western blot	ChIP-seq	RNA-seq
wt*)	+	+	y	y	wt1 (1487-1490) wt2 (1491-1494) wt3 (2192-2195)	wt1 (1740) wt2 (1741)
E469-2	-	+	y	y	1 (1301-1304) 2 (2196-2199)	(1736)
E469-3	-	+	y	y	1 (1305-1308) 2 (2200-2203)	1 (1737)
E471-2	+	-	y	y	n	n
E480-2	+	-	y	y	(1309-1313)	(1738)
E480-6	+	-	y	y	(1314-1318)	(1739)
dn*)	-	-	y	y	dn1 (1495-1498) dn2 (1499-1502) dn3 (2188-2191)	dn1 (1742)
E124-1	-	-	n	y	(1624-1627)	(1743)
E124-3	-	-	n	y	(1628-1631)	(1744)

+ and – indicate whether or not the gene is present and expressed in each cell line. Yes (y) or no (n) indicate whether or not the microscopy experiment (with immunofluorescence staining for H3K9me3 and HP1α), Western blot (detecting H3K9me3), ChIP-seq (using an H3K9me3-specific antibody) or RNA-seq have been performed. The numbers in parentheses correspond to the Hi-seq database ID. Some ChIP-seq samples were excluded due to quality issues (see below). *) original cell lines from (Peters *et al.* 2001). **) This sample was included in the PCH analysis (**Fig. 17**), but excluded from all other analyses since it did not have an IgG control for normalization.

The input samples of chromatin immunoprecipitation followed by DNA sequencing (ChIP-seq) confirmed the homologous recombination at SUV39H loci since no reads mapped at the respective positions within the SUV39H1 and SUV39H2 genes (**Fig. 13a**).

The number of reads that mapped to the respective SUV39H isoforms from strand-specific RNA sequencing (RNA-seq) was significantly reduced in the null cell lines confirming the functional knock-out (**Fig. 13b**). Note that the remaining low transcript levels might stem from regions upstream of the knock in site or spurious transcription through the *lacZ* gene, but are not expected to yield functional protein products.

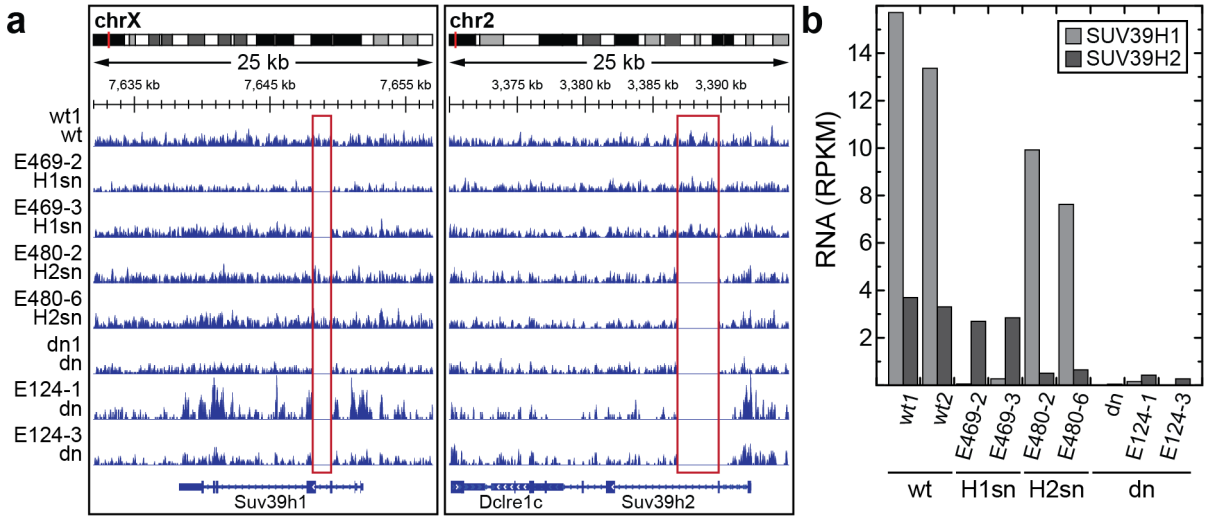

Figure 13: Confirmation of functional knock-out of SUV39H isoforms in iMEF cell lines. (**a**) Integrative Genomics Viewer (IGV, (Robinson *et al.* 2011, Thorvaldsdottir *et al.* 2013)) traces of ChIP-seq input samples. No reads mapped to a ~1.4 kb region within the *Suv39h1* gene and a ~3.0 kb region within the *Suv39h2* gene in the respective cell lines confirming disruption of the gene loci by knock-in of *lacZ*. (**b**) SUV39H1 (NM_011514) and SUV39H2 RNA (NM_022724) levels from strand-specific RNA-seq of the different iMEF cell lines. As expected, expression was significantly reduced in the respective null cell lines.

II.2.2 Antibody validation and ChIP-seq quality control

A number of experiments conducted in this study, namely immunofluorescence staining, western blot and ChIP-seq, depend on the use of H3K9me3-specific antibodies. The specificities of two different antibodies (from Abcam and Active Motif, see **Table 9** in the Materials & Methods section) were tested using a modified histone peptide array (**Fig. 14a**). The most relevant modifications are highlighted (**Fig. 14**) and listed in **Supplementary Table S1**. Both antibodies bound H3K9me3 as expected (**Fig S1b,c**), but also H4K20me3 (Abcam antibody) or H3K27me3 (Active Motif antibody). Both antibodies did not recognize H3K9me3 in combination with serine 10 phosphorylation (S10P) or threonine 11 phosphorylation (T11P).

As a quality control for ChIP-seq experiments, the similarity of the replicates was evaluated by calculating the Pearson correlation of read counts within 10 kb windows across the entire genome (**Fig. 15a-e**) using deepTools2 (Ramirez *et al.* 2016). If an input sample did not correlate well with the others, the entire sample (i. e. all controls and the

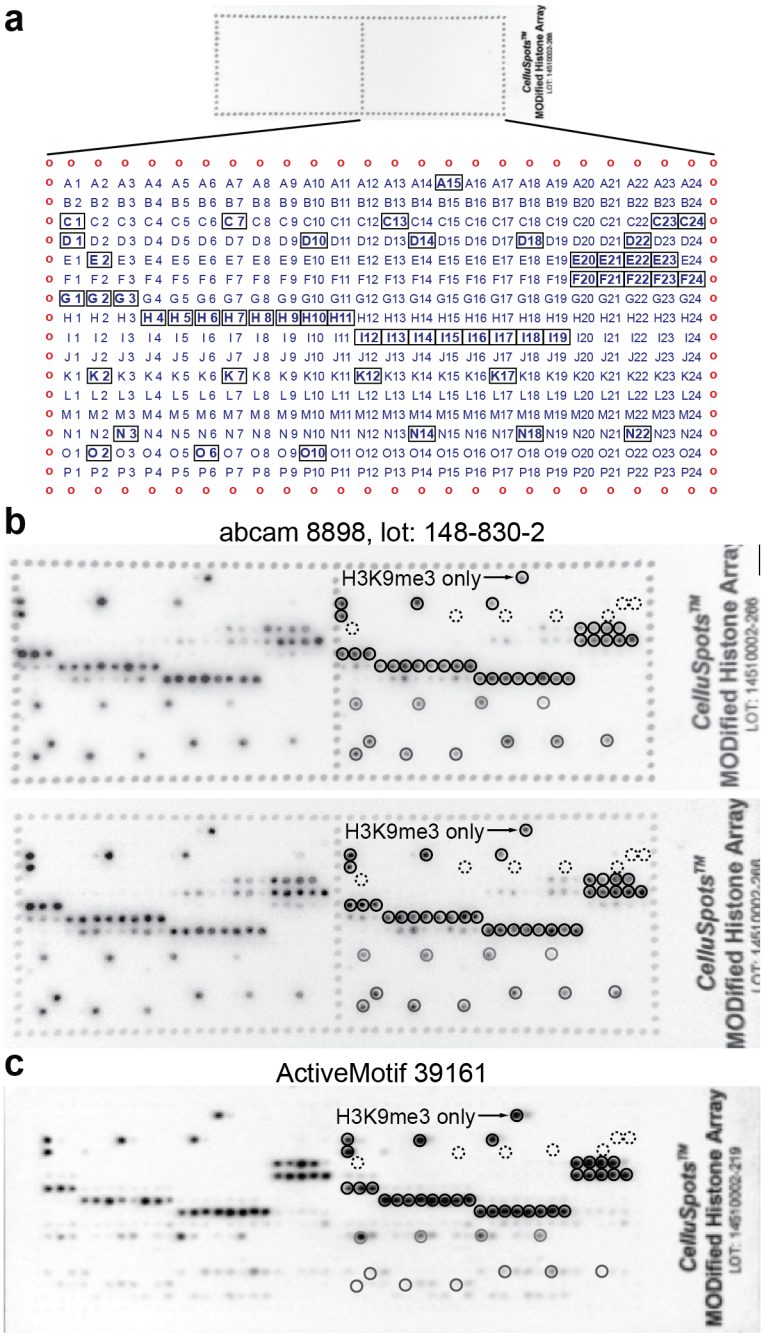

Figure 14: Modified histone peptide recognition by H3K9me3-specific antibodies. (**a**) Modified histone peptide array setup and selected positions marked in the images below. The modifications corresponding to the marked positions are listed in **Supplementary Table S1**. (**b**) The Abcam antibody recognized various H3K9me3-modified peptides (black), but not in combination with S10P or T11P (dashed). It also recognized some H3K27me3-modified peptides (light grey) as well as H4K20me3 (dark grey). Shown are two replicates where the antibody was tested on a new array (top) or on an array reused after stripping another antibody (bottom). (**c**) The Active Motif antibody bound similar peptides as the Abcam antibody (**b**), but bound weaker to H4K20me3-modified peptides and stronger to H3K27me3-modified peptides.

specific IP) was excluded from the analysis (e. g. wt3, see **Fig. 15**a). The correlations between controls and IP were very high for iMEF samples (e. g. for SUV39H2 cell lines, **Fig. 15d**). In contrast, the IPs of mESC samples correlated well with one another, but not with any of the controls (**Fig. 15e**), indicating that the IP signature was more dissimilar to the background signal (i. e. specific and/or enriched) in this cell type. The same was seen in the amplitude of the Pearson pseudo-autocorrelation at zero shift distance calculated by MCORE (Molitor *et al.* 2017) that was much higher for the mESCs (**Fig. 15f**).

For the remaining ChIP-seq samples, fingerprint plots were generated using deepTools2 (Ramirez *et al.* 2016) to check if the IP signal is enriched in a fraction of bins compared to the controls. The fingerprint plots (**Fig. 16**) are read as follows: 50% of the reads (first row, dashed horizontal line) fall into the ∼7% and ∼28% most covered bins (first row, right side of the dashed vertical line) for mESC and E469-2 ChIP, respectively. This indicates that there were fewer regions harboring a greater fraction of reads in mESCs, which points to a more peaked IP signature in mESCs and a much broader distribution of IP signal across many bins in iMEFs. In addition, the SUV39H dn samples in the last row did not show evenly distributed input patterns. In the most extreme case (E124-1), the 50% most covered bins (right side of the dashed vertical line) harbored >92% of the reads (above the dashed horizontal line), rendering the other 50% of the bins virtually not covered at all. Subsequently, features with insufficient input coverage were excluded from the read counting analyses (see below).

II.2.3 H3K9me3 levels in PCH and genome-wide depended mostly on SUV39H2

The wt, SUV39H sn and dn cell lines described above were used to systematically characterize different contributions of SUV39H1 and SUV39H2 to H3K9me3 both at PCH and elsewhere. Global levels of H3K9me3 were detected by western blot using the two different antibodies described above (**Fig. 14**). The SUV39H1 sn cell lines E469-2 and E469-3 had global H3K9me3 levels similar to the wt iMEFs, whereas the SUV39H dn cell lines E124-1 and E124-3 showed a strong overall reduction of H3K9me3 (**Fig. 17a**). This reduction of H3K9me3 was less pronounced, but nonetheless detectable, in the original dn cell line from (Peters *et al.* 2001) and in the SUV39H2 sn cell lines E471-2 and E480-6. With the exception of the SUV39H2 sn cell line E480-2, loss of SUV39H2 correlated with reduced global H3K9me3 levels.

To assess the role of the SUV39H enzymes at PCH, H3K9me3 and HP1α were stained by immunofluorescence and imaged by CLSM. PCH can be readily identified by intense DAPI staining and showed accumulation of H3K9me3 and HP1α in wt and SUV39H1 sn, but not in SUV39H dn iMEFs (**Fig. 17b,c**). SUV39H2 sn cells showed heterogeneous levels of enrichment which also varied for the different cell lines (Appendix, **Supplementary Fig. S1a**). Notably, the E480-2 cell line, which showed no global reduction of H3K9me3, did have reduced H3K9me3 levels in PCH compared to wild type cells, indicating that

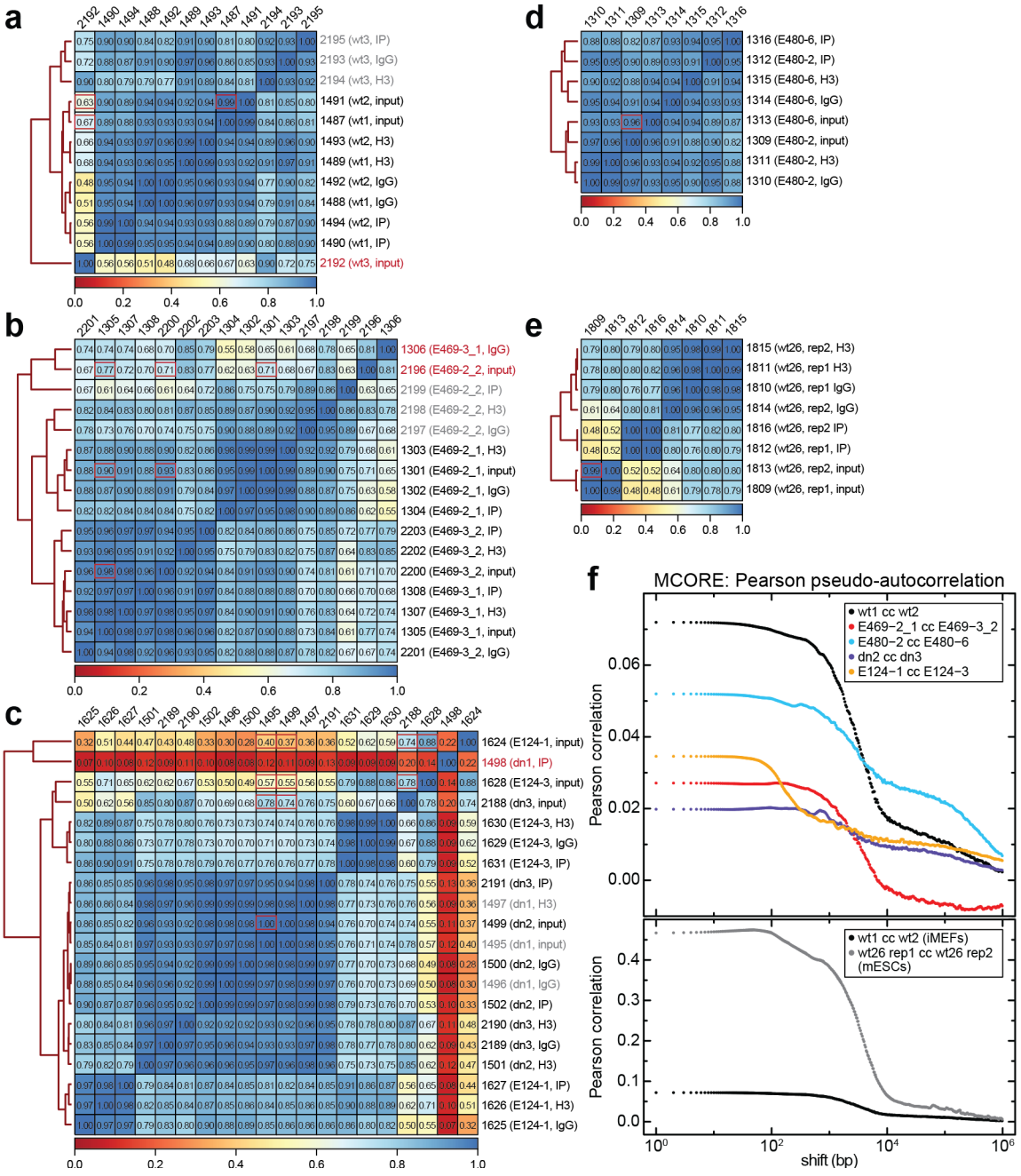

Figure 15: Pearson correlation as a quality control for ChIP-seq. In the heatmaps (**a**-**e**), red boxes indicate pair-wise correlations between input samples. Red and grey samples were excluded from the analyses. (**a**) Correlation of wt samples. The wt3 input did not correlate well with wt1 and wt2 inputs and wt3 was thus excluded. (**b**) Correlation of SUV39H1 sn samples. The E469-3_1 IgG control showed issues during sample preparation and was excluded. The E469-2_2 input correlated less with the other input samples than those among each other. E469-2_2 was thus excluded from the analyses. (**c**) Correlation of SUV39H dn samples. The dn1 sample was excluded since the IP did not work. Notably, E124-1 and E124-3 showed quite low pair-wise correlations with the dn cell line. In addition, the dn3 input did not correlate well with the dn1 and dn2 input. However, to have a dn cell line replicate, it was not excluded. (**d**) Correlation of SUV39H2 samples. (**e**) Correlation of two replicates of H3K9me3 ChIP-seq from mESCs for comparison. (**f**) Pearson pseudo-autocorrelation between replicates calculated by MCORE (Molitor *et al.* 2017). The amplitude at zero shift distance is a measure for replicate similarity. The input and IgG control were used for normalization and background correction, respectively. Top: iMEF cell lines. Bottom: wt iMEFs and mESCs.

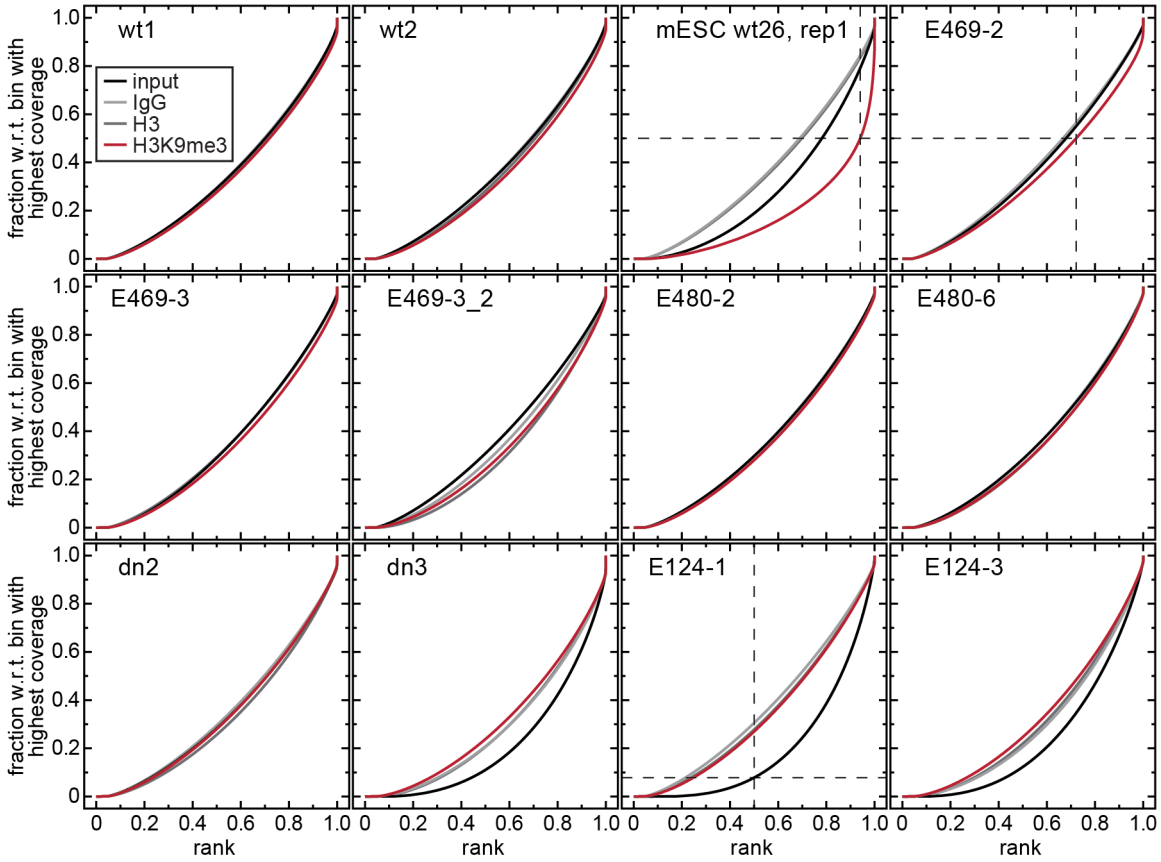

Figure 16: Fingerprint plots for all usable iMEF H3K9me3 ChIP-seq samples as well as an exemplary mESC H3K9me3 ChIP-seq data set for comparison. Fingerprints were generated using deepTools2 (Ramirez *et al.* 2016) based on (Diaz *et al.* 2012). The cumulative read counts are plotted against the rank of the bins (500,000 bins of 500 bp length) sorted according to reads that overlap with them. In iMEFs, most IP traces (red) were very similar to the controls as expected for broad distributions or non-specific IP.

H3K9me3 levels in PCH do not necessarily follow the global level. To verify the loss of H3K9me3 at PCH and to assess whether other genomic regions were also affected, ChIP-seq was carried out as described in the Methods section. A larger fraction of reads from ChIP-seq with an H3K9me3-specific antibody (Abcam) mapped to the consensus sequence of mouse major satellite repeats (Lehnertz *et al.* 2003) compared to input, non-specific ChIP with an IgG antibody or H3-specific ChIP (**Fig. 17d**). Similar to the microscopy analyses, different ChIP-seq read enrichments were observed among the various cell lines with the same genotype (Appendix, **Supplementary Fig. S1b**). In summary, H3K9me3 levels in PCH were reduced in cell lines lacking SUV39H2. Loss of SUV39H1 alone had little or no effect, but resulted in complete loss of the modification at PCH in combination with loss of SUV39H2. Together, these results indicate that SUV39H enzymes have an additive rather than an epistatic effect and that SUV39H2 contributes more. In addition, H3K9me3 levels in PCH did not always correlate with global levels, in particular among the SUV39H dn and the SUV39H2 sn cell lines.

II.2.4 Loss of pericentromeric H3K9me3 did not always lead to increased major satellite transcription

The effect of reduction or loss of H3K9me3 at PCH on transcription output was assessed by quantitative real-time PCR (qRT-PCR) and RNA-seq. In wild type and SUV39H1 sn cell lines, gamma satellite (GSAT) transcript levels were low (**Fig. 18**) as expected based on their high H3K9me3 levels at PCH (**Fig. 17** & **Supplementary Fig. S1**). GSAT transcription was upregulated in all SUV39H2 sn cell lines as well as in the original SUV39H dn cell line. Surprisingly, transcription was not significantly upregulated in the other two SUV39H dn cell lines (E124-1 and E124-3, **Fig. 18**) although they showed substantial global reduction of H3K9me3 and loss of H3K9me3 at PCH according to ChIP-seq (**Fig. 17a&d**). Conversely, the established MEF cell line NIH3T3, which is wild type for SUV39H, showed a substantial upregulation of GSAT transcripts despite having intact PCH in terms of H3K9me3, HP1, H4K20me3, MECP2, SUV39H and SUV4-20H (Muller-Ott *et al.* 2014). Taken together, transcriptional output from PCH was generally anticorrelated with H3K9me3, i. e. lower H3K9me3 resulted in higher RNA expression values. The wild type NIH3T3 cell line and the SUV39H dn cell lines E124-1 & E124-3 behaved differently. In the former, intact H3K9me3 did not prevent transcription. In the latter two cell lines, pronounced demethylation at H3K9 barely led to derepression of transcription. Those results indicate that H3K9me3 is generally associated with transcriptional repression, but other factors might interfere with this pathway.

II.2.5 SUV39H-dependent H3K9me3 outside of PCH

The enrichment of H3K9me3 at certain genomic features from ChIP-seq data was calculated taking into account non-specific background and input coverage as described in

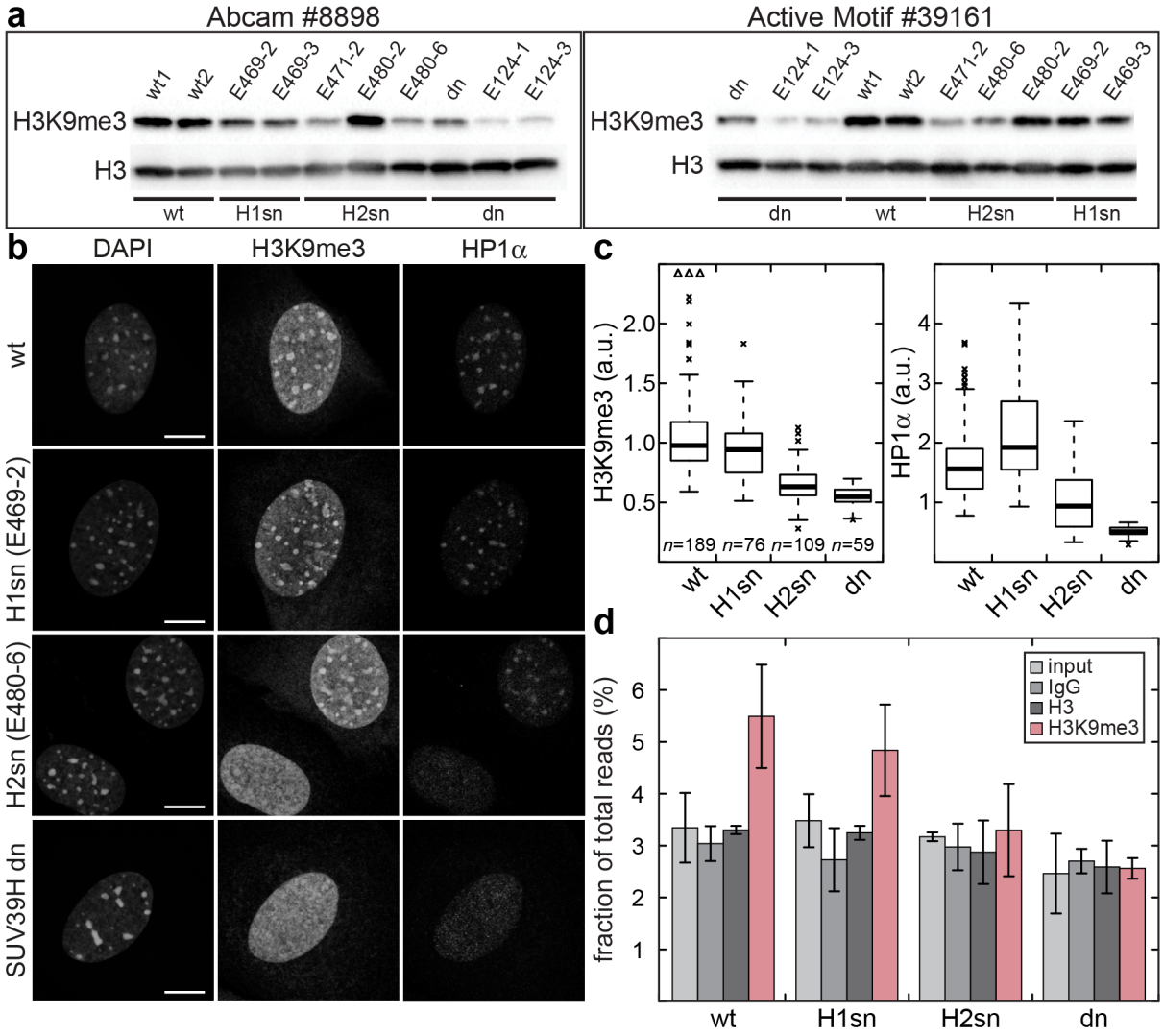

Figure 17: Global and PCH-specific levels of H3K9me3 in iMEF cell lines. (**a**) Western blots detecting H3K9me3 in histone extractions of iMEF cell lines. Two different H3K9me3-specific primary antibodies were used. Detection was based on chemoluminescence using horseradish peroxidase (HRP). (**b**) CLSM images of iMEF cell lines with immunofluorescent detection of H3K9me3 and HP1α. PCH was identified by dense DAPI staining (left). Scale bars: 5 µm. (**c**) Quantification of H3K9me3 and HP1α enrichment at PCH. Fluorescence intensity ratios were normalized to chromatin density. (**d**) Percentage of ChIP-seq reads mapping to the mouse major satellite consensus sequence. IP with an H3K9me3-specific antibody (Abcam) led to an enrichment of reads compared to input or IP with a non-specific antibody (IgG) or an H3 antibody. The corresponding single cell line plots for (**c** & **d**) are shown in the appendix (**Supplementary Fig. S1**).

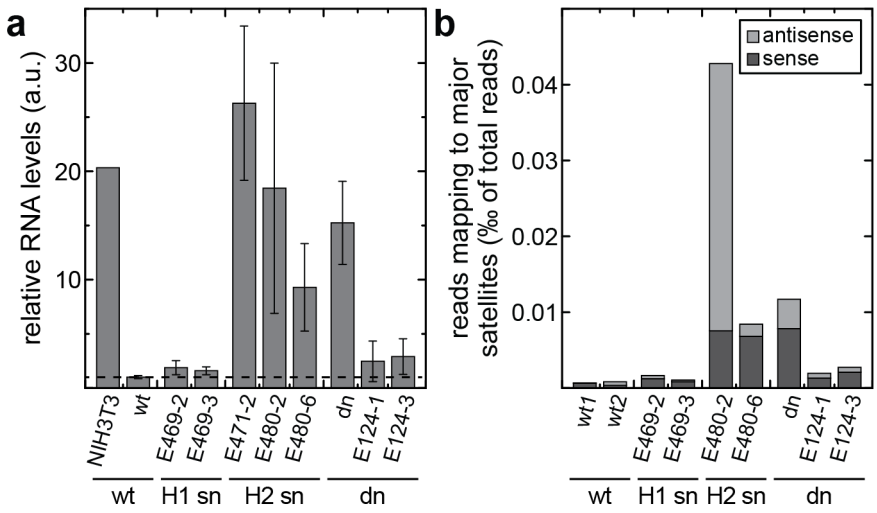

Figure 18: Major satellite transcript levels in different iMEF cell lines. (**a**) GSAT RNA levels determined by qRT-PCR. For each of three independent experiments (except NIH3T3 (n=1), E471-2 (n=2) and dn (n=5)), actin-normalized quantities were normalized to average wt levels. Error bars represent the standard error of the mean (sem) of the biological replicates. (**b**) Permillage of total reads from stranded RNA-seq that mapped to GSATs in sense and antisense direction. Notably, the results corresponded well to the qRT-PCR results shown in (**a**).

the Materials & Methods section (**Fig. 47**). As a reference, a mouse embryonic stem cell (mESC) H3K9me3 ChIP-seq experiment was used, because mESCs have been described to silence repetitive elements by SUV39H-dependent H3K9me3 (Bulut-Karslioglu *et al.* 2014). The reads mapping to the GSAT consensus sequence (**Fig. 19d**) were removed before mapping to the mm9 assembly of the mouse reference genome (see Materials & Methods, **Fig. 45** for details). There are 42 intergenic GSATs (Bulut-Karslioglu *et al.* 2012), which have single nucleotide polymorphisms compared to the consensus sequence and could thus be analyzed independently. 23 of the 42 intergenic repeats were mappable and 22 of those were highly methylated (i. e. >3-fold enrichment above genomic average) in wt iMEFs. Intergenic GSAT methylation in cell lines lacking either one of the SUV39H isoforms was reduced, although this effect was more pronounced for cell lines lacking SUV39H2 (**Fig. 19a**). The same analysis was done for 6387 repeat-rich (mainly LINEs and LTRs) genomic regions (HET6387) that have been described to be enriched for SUV39H-dependent H3K9me3 in mESCs (Bulut-Karslioglu *et al.* 2012). 5999 of those regions were mappable and most (5221 or ∼87%) were in fact highly trimethylated at H3K9 in mESCs. Only 319 (∼5%) of these were enriched for H3K9me3 in wt iMEFs (**Fig. 19b**). In contrast to GSATs, loss of SUV39H1 had only a very small effect, whereas SUV39H2 had again a greater effect on H3K9me3. The density plot for the H3K9me3 levels of these 319 regions (**Fig. 19b**, right) showed a single shifted peak for the SUV39H2 sn and SUV39H dn cell lines, indicating that the reduction of methylation was similar for

all selected regions.

179 RefSeq genes were highly methylated in wt iMEFs and 119 (∼66%) of those were protein-coding (NM IDs in the RefSeq annotation), whereas the other 60 encoded non-coding RNAs (NR IDs in the RefSeq annotation). In contrast to the HET6387 regions (**Fig. 19b**), methylation reduction in SUV39H sn/dn cell lines appeared to be bimodal for the RefSeq genes (**Fig. 19c**). This observation suggests that one population of genes is methylated by SUV39H and one population is not. As expected, nearly all genes that were affected by loss of one of the SUV39H isoforms were also affected by loss of both isoforms (**Fig. 19d**), and genes that were not affected in SUV39H dn cells were also not affected in the sn cell lines (genes in the right peak of the density plot that is centered close to 1). Separate analysis of protein-coding and non-coding RNA genes revealed that the bimodality was much more pronounced in the protein-coding genes (**Fig. 20**). Among the 119 H3K9-trimethylated protein-coding genes (**Supplementary Table S3**) were 26 genes coding for Zinc finger proteins (Zfp). Genes from this protein family have previously been found to be enriched in a subgroup of genes bound by HP1β (Vogel *et al.* 2006). However, all 26 Zfps were found in the population that did not change the H3K9me3 level (**Fig. 19c**) in all cell lines and were thus not strongly affected by loss of SUV39H enzymes.

Most repetitive sequences analyzed here showed average H3K9me3 levels slightly above the genomic average, but only GSATs and a subset of other heterochromatic regions showed high methylation levels (**Fig. 19**). Similar to PCH, those highly methylated regions were more affected by the loss of SUV39H2 than by the loss of SUV39H1. RefSeq genes displayed average H3K9me3 levels below the genomic average, but a small subset of genes was highly methylated in wt iMEFs and this high H3K9me3 level was SUV39H-dependent in some of the genes. Notably, more genes were affected by the loss of SUV39H2 than by the loss of SUV39H1 supporting the hypothesis that SUV39H2 plays a greater role in controlling H3K9me3 levels.

II.2.6 Upregulation of transcription upon loss of SUV39H-dependent H3K9me3 was not a universal feature across the genome

To assess the impact of SUV39H-dependent loss of H3K9me3 at the features analyzed above (**Fig. 19**), RNA was isolated and sequenced in wt, sn and dn iMEFs. Since the RNA isolation protocol included a selection step for long RNAs (>200 nt, see Materials & Methods section and **Supplementary Fig. S2**), GSATs shorter than 200 nt were excluded from the analysis. In wt iMEFs, methylated GSATs showed generally very low transcript levels (RPKM, **Fig. 21a**). In the SUV39H-deficient cell lines, only a few features showed slightly higher transcript levels and those appeared rather independent of the methylation decrease. The same was true for the methylated HET6387 regions (**Fig. 21b**). To evaluate the relationship of H3K9me3 and transcription at genes, the

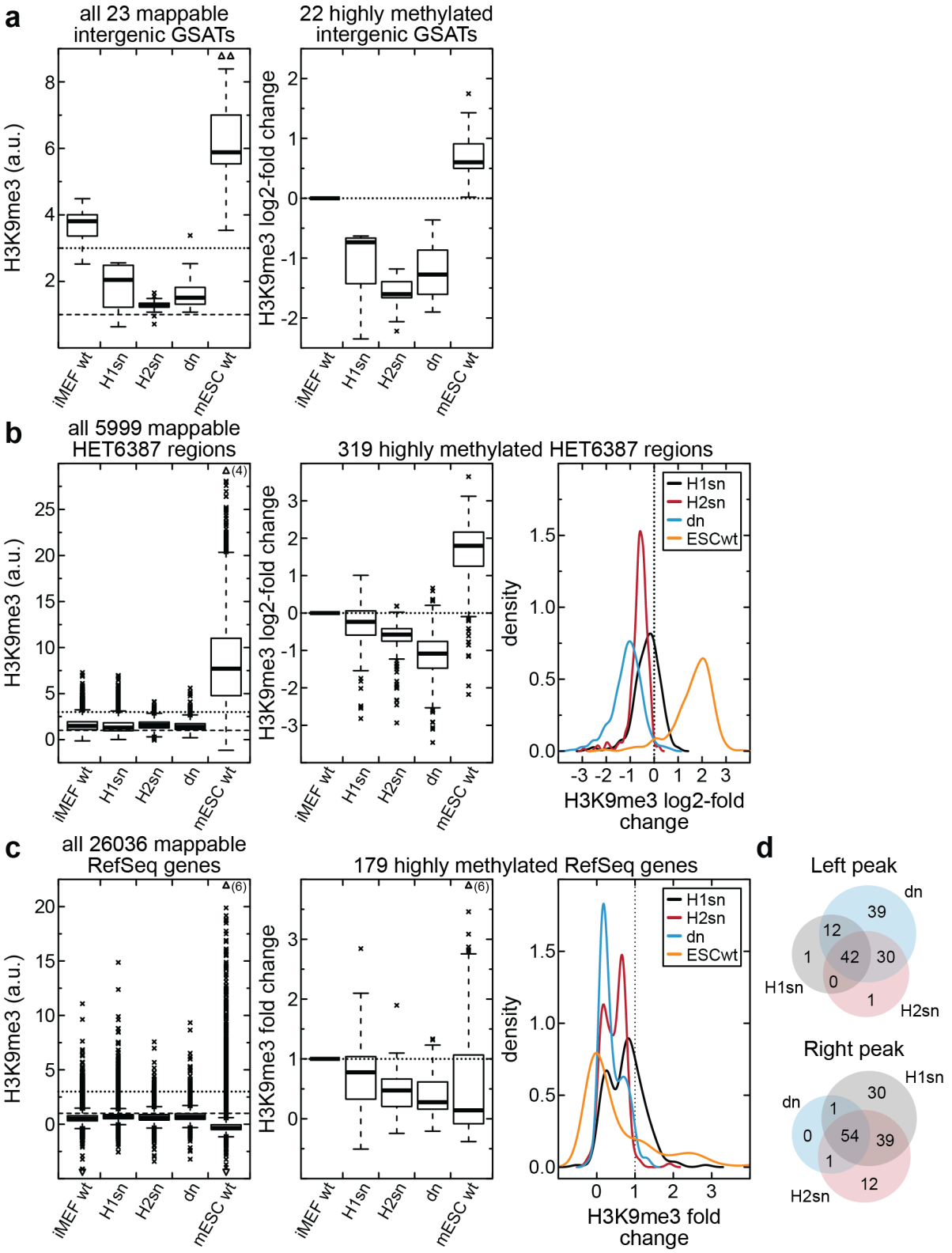

Figure 19: SUV39H dependency of H3K9 trimethylated features throughout the genome. In the left column, all features of the respective group are shown. Dashed line: genomic average. Dotted line: high methylation threshold (>3-fold enrichment above genomic average). In the center column, the log2-fold change of the highly methylated features (selected in wt iMEFs) normalized to wt iMEFs is shown. Dotted line: wt iMEF level. Where applicable, density plots of the center column are shown on the right. (**a**) Intergenic GSATs. (**b**) Regions displaying SUV39H-dependent H3K9me3 in mESCs (HET6387). (**c**) RefSeq genes. (**d**) Venn diagrams of genes in the left and right peaks of the density plot. The circle sizes are proportional to the number of genes in the respective peaks.

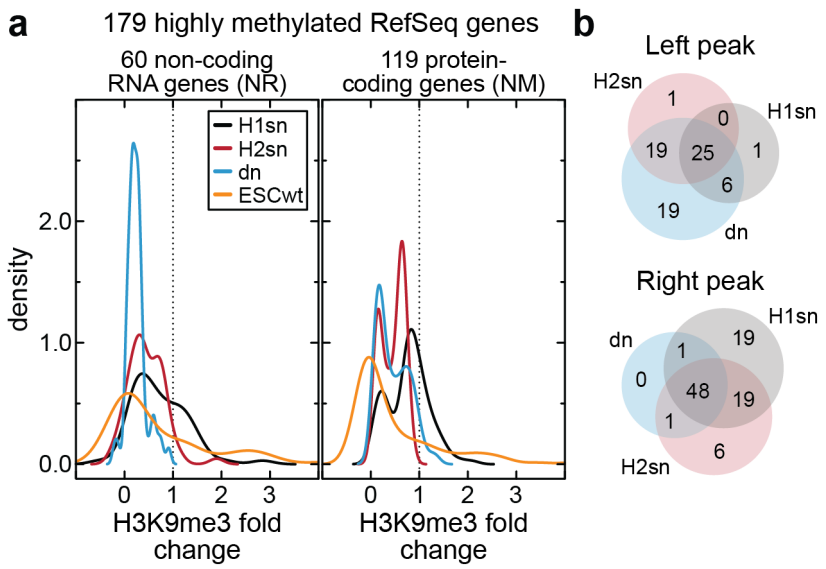

Figure 20: H3K9me3 fold change at non-coding and coding genes. (**a**) Density plots of H3K9me3 fold change of in wt iMEFs highly methylated non-coding (left) and protein-coding (right) RefSeq genes. The bimodality in **Figure 19c** mostly originated from the protein-coding genes (right, NM IDs in the RefSeq annotation). (**b**) Venn diagram of the 119 protein-coding genes in the left and right peak (same as in **Fig. 19d**, but without the non-coding genes).

long non-coding RNAs were excluded from the analyses below, since half of them were shorter than 200 nt and were thus eliminated during the RNA-seq library preparation. In addition, they displayed no clear bimodality when analyzed separately (**Fig. 20**). The 119 protein-coding genes that were highly methylated in wt iMEFs showed very heterogeneous absolute transcript levels (**Fig. 21c**, orange dots). Some of the strongly demethylated genes (i. e. log2-fold change of around -3) showed very high transcript levels (**Fig. 21c**, upper left corner). To assess the change in transcript levels upon loss of SUV39H, the log2-fold change of transcripts as calculated by DESeq2 (Love *et al.* 2014) was plotted against the change in H3K9me3 (**Fig. 21d**). In general, genes that showed demethylation upon loss of SUV39H showed either upregulation or no change in transcription (**Supplementary TableS3**), yielding a slight anticorrelation between H3K9me3 and RNA changes. The 26 Zfp genes that were highly methylated in wt iMEFs showed moderate expression levels in wt iMEFs (3.1±2.9 RPKM, mean & sd) indicating that they were not silenced despite being heavily trimethylated at H3K9 (**Fig. 21e**). Those expression levels were mostly unaffected by SUV39H loss (**Fig. 21f**) in agreement with the result above that their methylation level is not strongly affected by the loss of SUV39H enzymes (**Fig. 19c&d**).

119 protein-coding genes were highly methylated in wt iMEFs (**Fig. 19** and **Fig. 20**) with the 2900026A02Rik gene (NM_172884) displaying the highest change in RNA expression (5.6 – 7.0 in the different cell lines, see **Fig 21d**, dashed circle). Interestingly, the high H3K9me3 level in wt iMEFs (∼4-fold enriched above genomic average) was re-

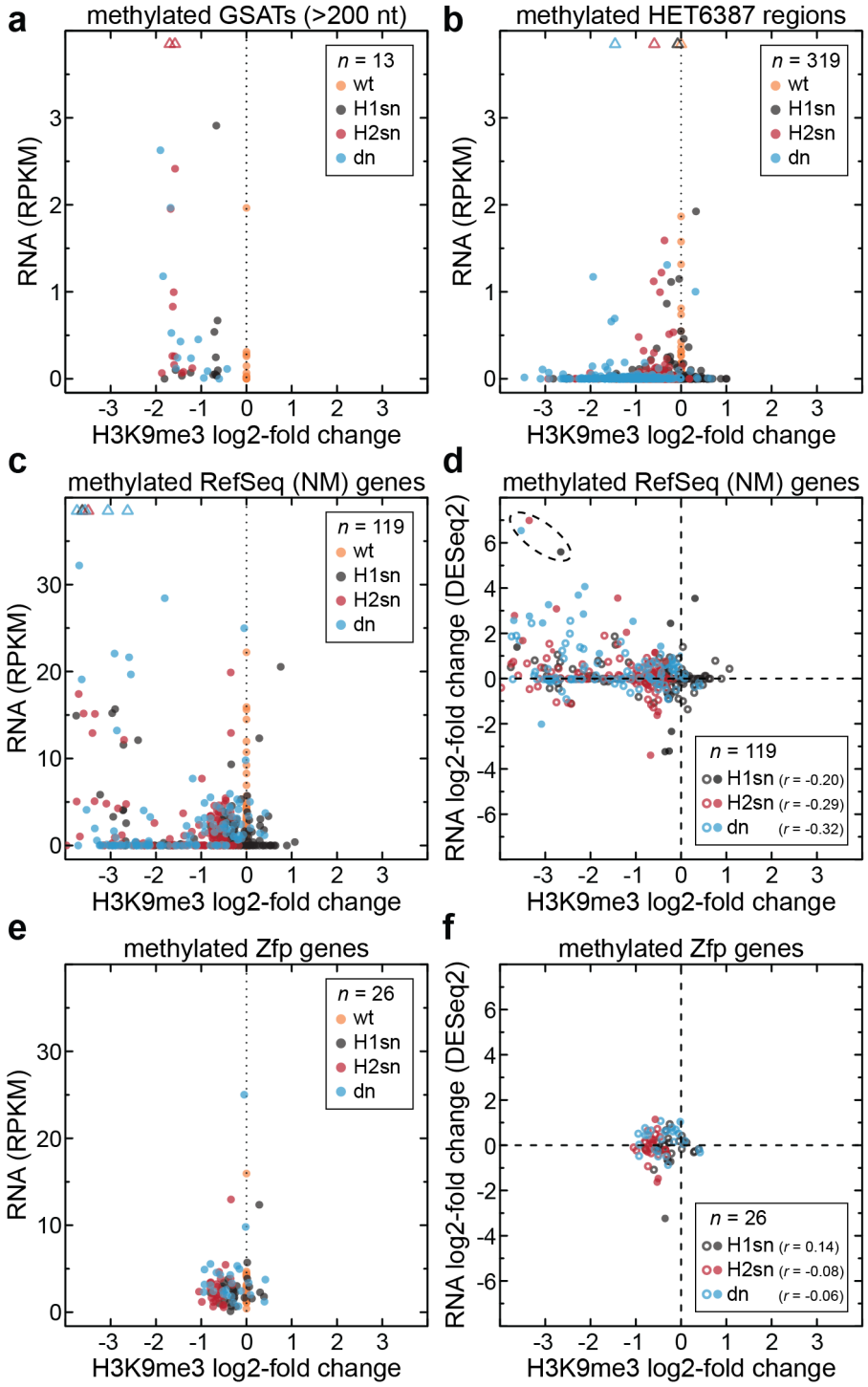

Figure 21: Transcript levels of in wt iMEFs methylated genomic regions as determined by RNA-seq. Absolute read densities (reads per kilobase of exon per million mapped reads, RPKM) were plotted against the log2-fold change in H3K9me3 (relative to wt iMEF levels, **a-c**,**e**). Each dot represents one genomic region and the number of regions are given in the legends. (**a**) Methylated intergenic GSATs. (**b**) Methylated HET6387 regions. (**c**) Methylated protein-coding RefSeq genes. (**d**) The RNA log2-fold change (compared to wt iMEFs) determined by DESeq2 was slightly anticorrelated with the H3K9me3 log2-fold change (the Pearson correlation coefficient r is given in the legend). Filled circles represent significant RNA changes (DESeq2 output $p_{\text{adj}} < 0.05$) and the points marked by the dashed ellipse correspond to the 2900026A02Rik gene (NM_172884) shown in **Figure 22a**. (**e&f**) Same analyses as in (**c&d**), but only for the subset of in wt iMEFs methylated Zfp genes.

duced to $\leq 16\%$ of the wt level in the functional knock out cell lines (**Supplementary Table S3**). Visualizing the H3K9me3 signature revealed that this gene was located in a large SUV39H-dependent H3K9me3 domain (**Fig. 22a**). However, loss or reduction of this domain did not lead to an upregulation of all genes within this domain, suggesting that demethylation of H3K9 is not sufficient to activate gene expression. Complementary to the approach above (**Fig. 21c–f**), where all genes with high at H3K9me3 levels in wt iMEFs were analyzed with respect to their methylation and expression changes, all differentially expressed genes were also analyzed with respect to their H3K9me3 levels. Using rather stringent parameters to partition the DESeq2 results ($p_{\text{adj}} < 0.01$ and log2-fold change > 3 or < -3) resulted in a common set of 64 upregulated and 45 downregulated genes (**Fig. 22b**, **Supplementary Table S4**). The upregulated genes showed no trend for changing their H3K9me3 levels in any direction (**Fig. 22c**). Surprisingly, the strongly downregulated genes tended to show slightly increased H3K9me3 levels (**Fig. 22c**) despite having lost one or both isoforms of the SUV39H histone methyltransferase.

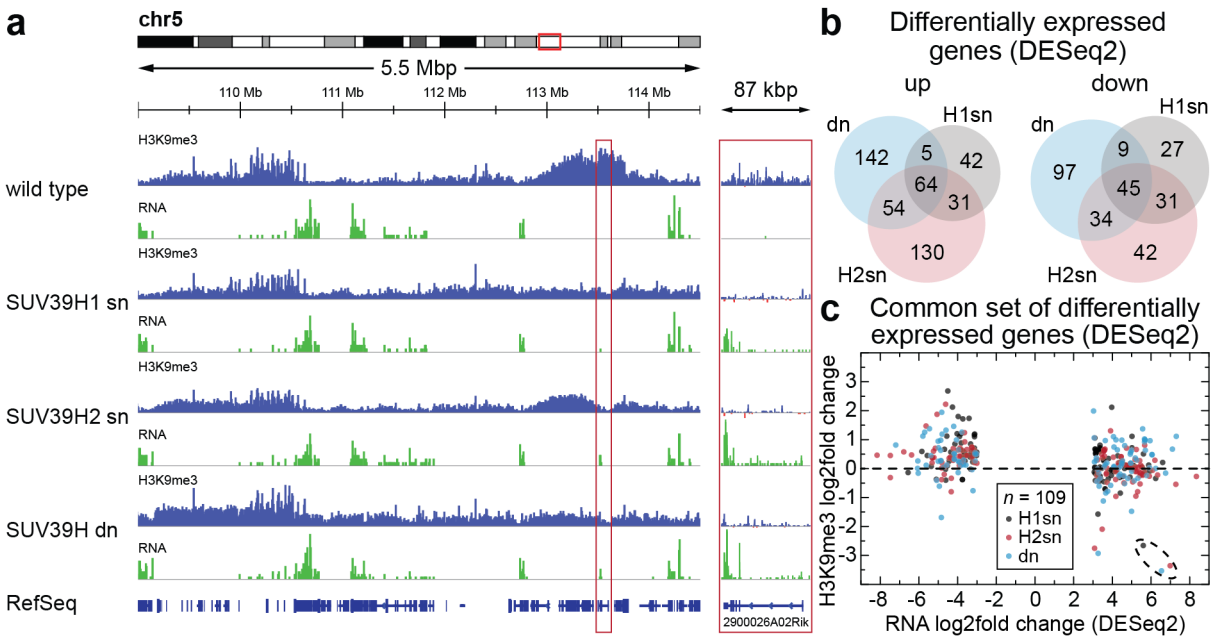

Figure 22: H3K9me3 at differentially expressed genes in different SUV39H functional knock out cell lines. (**a**) H3K9me3 ChIP- and RNA-seq signatures of a genomic region on chromosome 5. ChIP-seq traces were normalized with MCORE (Molitor *et al.* 2017) and single traces of cell lines with the same genotype were aggregated and visualized with IGV (Robinson *et al.* 2011, Thorvaldsdottir *et al.* 2013). The RNA-seq trace was scaled to account for the number of mapped reads in each sample. The right red box is an enlargement of the left red box showing the 2900026A02Rik gene (NM_172884). (**b**) Differentially expressed genes as determined by DESeq2 (Love *et al.* 2014). Gene expression with a log2-fold up- (left) or down-regulation (left) >3 and an adjusted p-value <0.01 were considered to be significant. (**c**) H3K9me3 log2-fold change of the common set of differentially expressed genes (i. e. 64 up- and 45 downregulated genes). The dashed line represents the wild type methylation level and the 290026A02Rik gene shown in panel (**a**) is marked by a dashed circle.

In summary, SUV39H-dependent loss of H3K9me3 had little to no effect on the transcription of intergenic GSATs and heterochromatic regions (**Fig. 21a&b**) that have previously been shown to be SUV39H-dependent in mESCs (Bulut-Karslioglu *et al.* 2012). At PCH (**Fig. 18**) and at genes that are highly methylated in wt iMEFs (**Fig. 21d**), transcription was either upregulated or did not change (compared to wt levels) indicating that SUV39H-dependent loss of H3K9me3 was not sufficient to activate transcription in all cases, but rendered some loci more permissive towards activation. The fact that some highly methylated genes in wt iMEFs and intact PCH in the NIH3T3 cell line displayed high levels of transcripts (**Fig. 21c** & **Fig. 18a**) indicates that H3K9me3 was not sufficient to completely silence genomic loci.

II.3 Activating transcription of a heterochromatic gene array with light

Heterochromatic loci can transition from a repressed chromatin state to an activated one when stimulated, e. g. by environmental cues. In addition, some active genes can be located within heterochromatin domains indicating that repressed compartments might not be completely silenced (Dimitri *et al.* 2005). Thus, understanding the activation of a repressed locus might help to decipher the different mechanisms guiding appropriate and inappropriate transitions from silent to active chromatin states. To dissect the events at a silenced locus upon activation of transcription, an optogenetic transcription activation system was established. Therefore, the transcription reporter cell line U2OS 2-6-3 (Janicki *et al.* 2004) was combined with the optogenetic CIB1/CRY2 system (Kennedy *et al.* 2010).

II.3.1 Effector proteins were recruited to subnuclear compartments by blue light

To quickly recruit effector proteins such as transcription activators to subnuclear compartments, I developed an optogenetic system termed Blue Light-Induced Chromatin Recruitment (BLInCR). To this end, effectors were fused to a fluorescent protein (FP) and the PHR domain of the *Arabidopsis thaliana* protein CRY2 (Kennedy *et al.* 2010). Its interacting partner, the CIBN domain of CIB1, was fused to an FP-tagged 'localizer', i. e. a protein with a specific subnuclear localization (**Fig. 23a**). PHR-mCherry was used as a mock effector to test the versatility of the system. Before illumination with blue light, PHR-mCherry was rather evenly distributed throughout the cells and depleted from nucleoli (**Fig. 23**, center column). The CIBN domain was fused to a variety of proteins (see **Table 6** in the Materials & Methods section) thereby targeting different compartments within the nucleus: a repetitive array stably integrated into the genome, nucleoli, telomeres, nuclear lamina and PML nuclear bodies (**Fig. 23b**, left column and schematic drawings from top to bottom). Upon blue light exposure, the PHR-tagged effector relocalized to sites of CIBN (**Fig. 23b**, right column). The BLInCR system could also be used in other systems such as PCH foci in wt iMEFs. Here, murine SUV39H1, SUV4-20H1 and SUV4-20H2 were used as 'localizers' and tagged with CIBN-YFP. PHR-mCherry-NLS served as a mock effector and could be recruited as expected (**Fig. 24a**). To verify the localization of the 'localizer' proteins when tagged with CIBN, cells transfected with CIBN-mSUV39H1 and a PHR-mCherry-effector construct were fixed and counterstained with DAPI after blue light exposure. As expected, the recruited constructs colocalized with DAPI-dense foci (**Fig. 24b**) confirming recruitment of the effector to PCH. In summary, all 'localizer' proteins tested thus far were compatible with the BLInCR system.

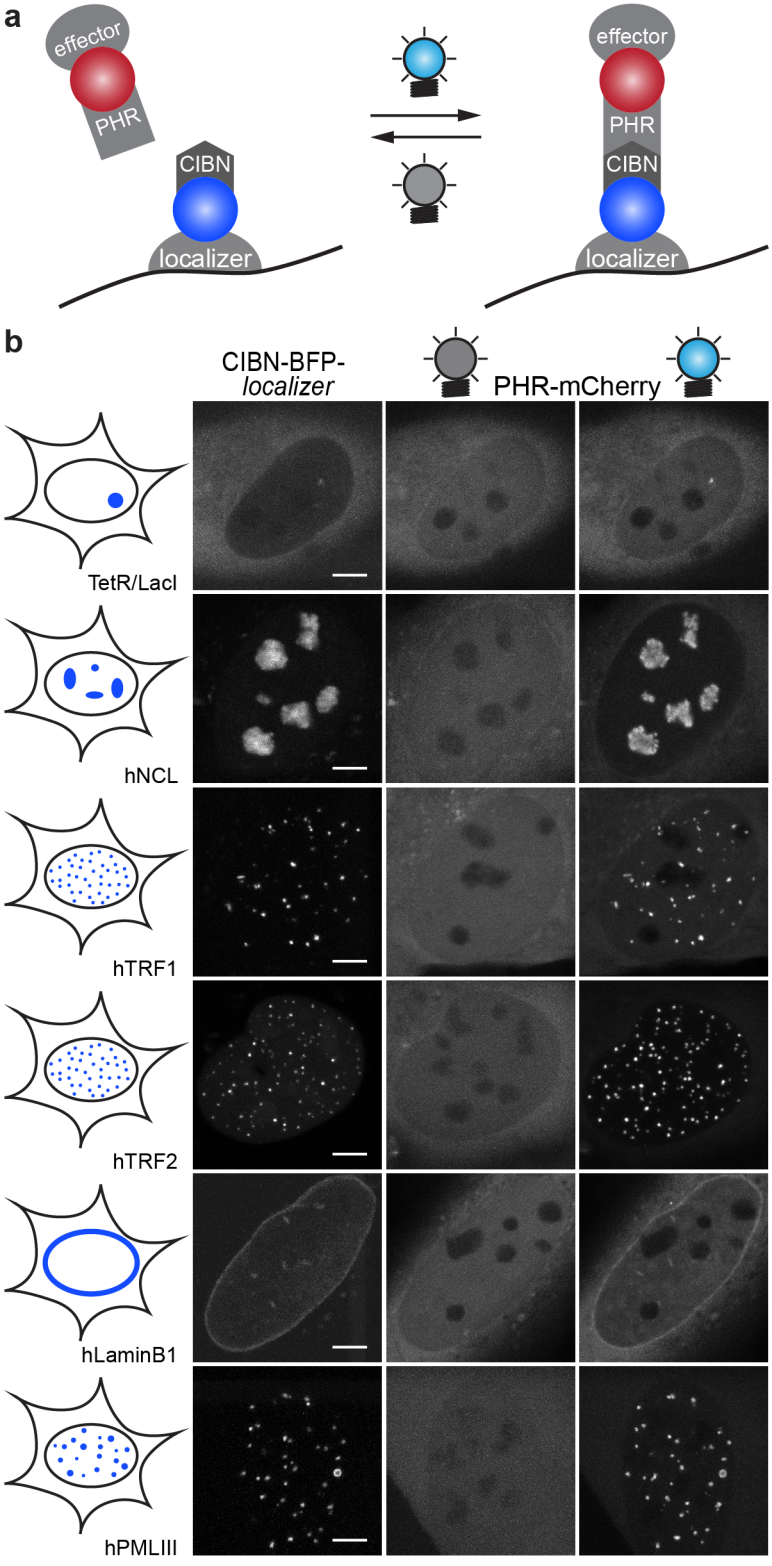

Figure 23: Blue Light-Induced Recruitment to Chromatin (BLInCR) in U2OS 2-6-3 cells. (**a**) Principle of BLInCR: the CIBN domain is tethered to a 'localizer' and a PHR-tagged effector protein relocalizes to the site of the 'localizer' upon exposure to blue light. (**b**) Different 'localizer' proteins were tagged with CIBN and tagBFP for visualization (left). PHR-mCherry was evenly distributed throughout cell (center) and relocalized to sites of the 'localizers' after illumination with blue light. All constructs were transiently transfected. Scale bars: 5 µm. Figure adapted and modified from (Rademacher *et al.* 2017).

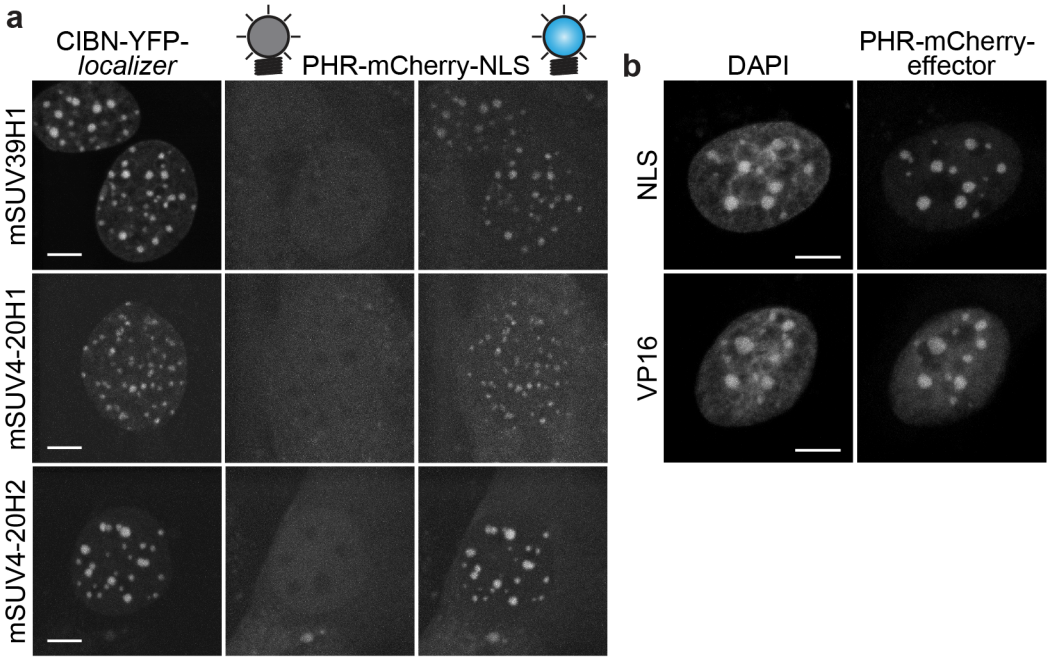

Figure 24: Adaptation of BLInCR to PCH in MEF cell lines. (**a**) Different PCH components were used as 'localizers' and tagged with CIBN-YFP: murine SUV39H1, SUV4-20H1 and SUV4-20H2 (from top to bottom). PHR-mCherry-NLS colocalized with the CIBN counterparts after illumination with blue light (in wt iMEFs). Scale bars: 5 µm. (**b**) NIH3T3 cells were transfected with CIBN-mSUV39H1 (not visible) and exposed to blue light before fixation. PHR-mCherry-effector constructs (NLS or VP16) colocalized with DAPI dense foci indicating that recruitment via CIBN-mSUV39H1 worked as expected. Scale bars: 5 µm.

II.3.2 BLInCR was rapid and reversible

The human osteosarcoma cell line U2OS 2-6-3 contains a stably integrated construct with promoter-proximal binding sites for LacI and TetR and a reporter cassette comprising a gene coding for peroxisome-localized CFP as well as MS2 loop sequences (**Fig. 25a**). Around 200 copies of this construct are integrated near the telomere of the short arm of chromosome 1 (Janicki *et al.* 2004). This cell line was used to characterize the binding and dissociation kinetics of PHR-tagged effectors to array-tethered CIBN.

For recruitment, the array was located without triggering the switching of the PHR domain (excitation with a wavelength of $\lambda = 561$ nm). Subsequently, the PHR-YFP-tagged effector was imaged, thereby inducing PHR switching and binding to CIBN ($\lambda = 488$ nm). After <2 s, accumulation of fluorescence intensity was already visible (**Fig. 25b**). For the quantification, a circular region around the array and a reference region were selected in the maximum intensity projection of the time series (see Materials & Methods section for details). PHR-YFP tagged effectors accumulated rapidly at the array with two characteristic rates and characteristic times to reach half-maximal levels in the order of 10-23 s **(Table 4**). The PHR-YFP-fused nuclear localization signal (NLS) displayed slightly slower accumulation compared to GCN5 and VP16 (**Fig. 25c**). This was not an artifact of the NLS since PHR-YFP showed very similar recruitment kinetics and the same was true for PHR-YFP-hGCN5 and PHR-YFP-hGCN5-NLS (**Fig. 25d**, left, **Fig. 25e**). In addition, the recruitment kinetics to *tet*O or *lac*O were similar to each other but slightly faster for *lac*O, as shown for PHR-YFP-VP16 and PHR-YFP-hGCN5 (**Fig. 25d**, right, **Fig. 25e**). The recruitment curves were fitted to a model considering two parallel first-order reactions ($E(t) = a - b \cdot \exp(-k_1 \cdot t) - c \cdot \exp(-k_2 \cdot t)$). The two rates could correspond to PHR-CIBN heterodimerization and the PHR-PHR oligomerization that has previously been described (Bugaj *et al.* 2013). In this case, the rates should be identical for all PHR and CIBN constructs and only differ in their relative contribution to the overall recruitment kinetics (i. e. the fast and slow fraction b and c. Accordingly, all 123 curves were fitted with a single set of rates k_1 and k_2, but individual plateau values a and contributions of the fast and slow reaction b and c to account for different transfection efficiencies and the properties of the fused effectors. All curves and fits were normalized and averaged, resulting in the curves shown in **Figure 25c&d**. For comparison, single curves were fitted to the same model leaving all five parameters unconstrained. This was difficult because some of the recruitment curves had a low signal-to-noise ratio and the fit results frequently depended on the starting values. The normalization and averaging was done as before and the average curves were refitted. Results from both fitting strategies were similar (**Table 4**) and the characteristic recruitment times were heterogeneous among cells, which was more pronounced for the single fits (**Fig. 25e**). The fit quality was comparable for the two fit methods (**Fig. 25f**) and the cumulative sum of squared residuals from the global fit were only 9% higher than those of the single fits, indicating

that global fitting of the two characteristic rates was suitable as it described the data well and was more robust to noise and individual outliers.

Table 4: Fit parameters for BLInCR kinetics.

PHR-YFP fused to	**VP16**		**hGCN5**		**hGCN5-NLS**	**NLS**	-
recruited to	*tet*O	*lac*O	*tet*O	*lac*O	*tet*O	*tet*O	*tet*O
number of cells n	18	10	18	11	32	19	15
global fit of k_1 and k_2							
k_1 (s^{-1}), fast reaction [a]	0.127±0.002						
b, fraction fast [b]	0.41±0.17	0.47±0.19	0.40±0.23	0.53±0.26	0.36±0.17	0.19±0.21	0.23±0.24
k_2 (s^{-1}), fast reaction [a]	0.0230±0.0004						
c, fraction slow [b]	0.48±0.18	0.32±0.18	0.46±0.26	0.34±0.26	0.44±0.21	0.73±0.22	0.67±0.32
$\tau_{1/2}$ (s) [b]	13.6±5.0	10.8±4.5	14.1±7.1	11.1±6.6	14.0±5.6	22.2±8.0	20.6±9.3
single curve fits for normalization and refitting of average curves							
k_1 (s^{-1}), fast reaction [a]	0.30±0.03	0.30±0.03	0.73±0.21	1.94±1.98	0.38±0.06	0.71±0.72	0.29±0.12
b, fraction fast [b]	0.26±0.01	0.34±0.02	0.12±0.02	0.08±0.06	0.21±0.02	0.02±0.01	0.09±0.02
k_2 (s^{-1}), fast reaction [a]	0.039 ± 0.001	0.041 ± 0.002	0.049 ± 0.001	0.066 ± 0.002	0.041 ± 0.001	0.034 ± 0.001	0.033 ± 0.001
c, fraction slow [b]	0.64±0.01	0.48±0.02	0.71±0.01	0.76±0.01	0.61±0.01	0.78±0.004	0.73±0.01
$\tau_{1/2}$ (s) [*,b]	11.9±5.6	10.4±8.1	13.4±7.1	10.6±6.5	12.0±6.9	25.9±12.3	23.8±13.1

[a] These fit results are listed along with their standard fit error. [b] These fit results correspond to the average of the single fit results and are listed with their standard deviations. [*] $\tau_{1/2}$ is the characteristic time to reach half-maximal levels.

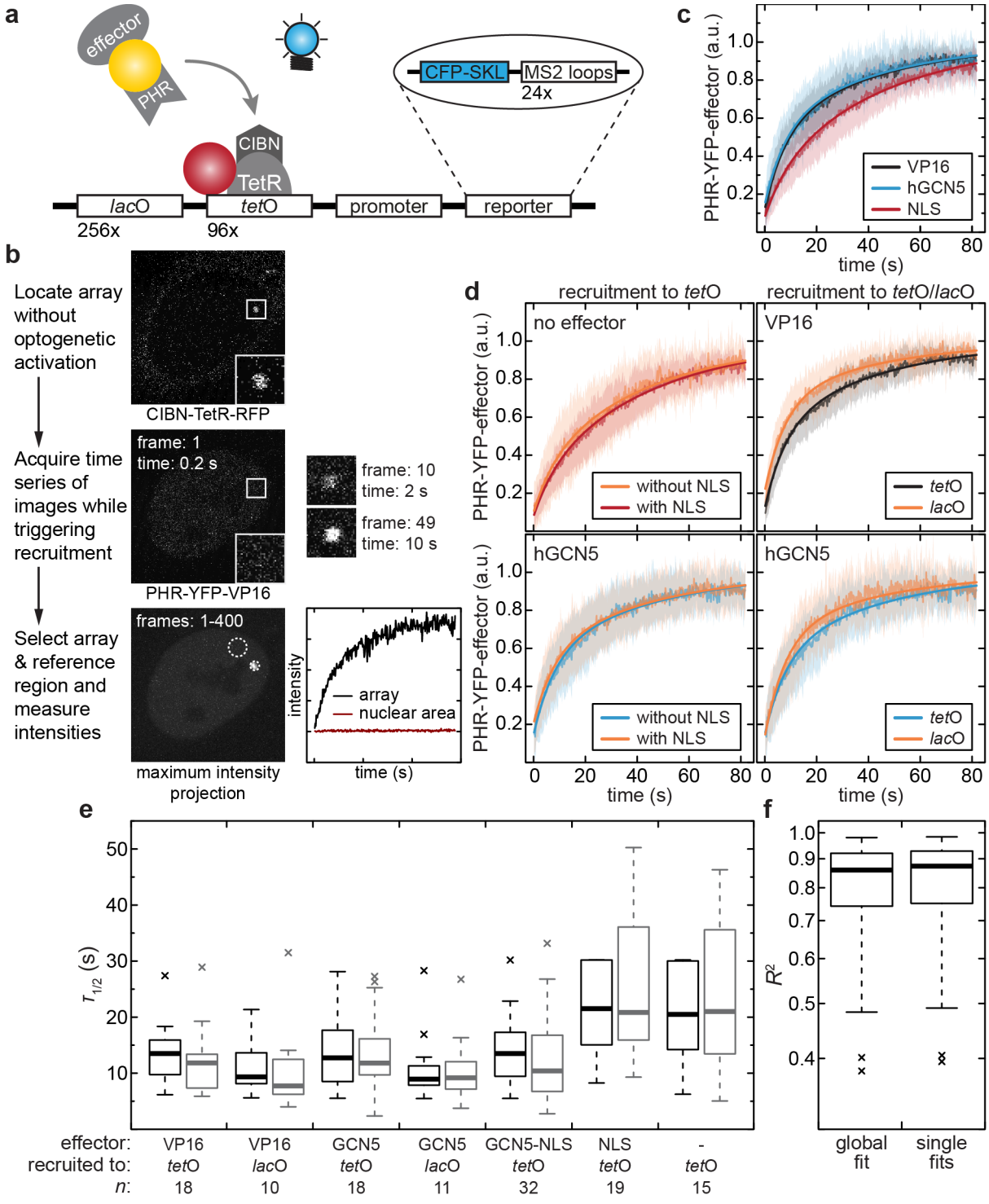

Figure 25: BLInCR kinetics. (**a**) Schematic representation of the array and recruitment setup. Around 200 copies of this construct are incorporated in the U2OS 2-6-3 cell line (Janicki *et al.* 2004). (**b**) Workflow for BLInCR kinetics measurements. See text for details. (**c**) Measured BLInCR kinetics for different effectors recruited to the *tet*O sites. (**d**) BLInCR kinetics of (mock) effector with and without NLS to *tet*O sites (left) and effector recruitment to *tet*O compared to *lac*O (right). The black, blue and red curves are identical to those in (**c**). (**c&d**) Averaged experimental data (transparent lines) with standard deviations (shaded areas) and averaged global fits (solid lines) are shown, fit parameters are listed in **Table 4**. (**e**) Characteristic recruitment times determined from the global (black) or the single curve fits (grey). (**f**) Coefficients of determination for both fit methods. Scaling of the *y*-axis is logarithmic. Parts of the figure are adapted and modified from (Rademacher *et al.* 2017).

To characterize the reversibility of BLInCR (**Fig. 26a**), the effectors were tagged with PHR-mCherry to allow for visualization without triggering recruitment ($\lambda = 561$ nm). After recording a pre-recruitment stack, the GFP-marked array was visualized ($\lambda = 488$ nm), thereby triggering light-induced PHR switching and recruitment of the effectors. Subsequently, imaging of the effector construct without recruiting anew allowed to track their dissociation from the array. As for the BLInCR kinetics measurement, the fluorescence intensity at the array and a nuclear reference area were measured (**Fig. 26b**, see Materials & Methods section for details). PHR-mCherry-fused effectors readily dissociated from the array (**Fig. 26c**) with a characteristic half-life time of ∼5 min (**Fig. 26d**, **Table 5**) which was identical for both effectors tested (VP16 and NLS). Importantly, the effectors completely dissociated from the array within 15-20 min. In addition, they could be rapidly recruited again as shown for NLS BLInCR to PCH via CIBN-EYFP-SUV4-20H2 (**Fig. 26e**).

Table 5: Fit parameters for BLInCR dissociation kinetics.

PHR-mCherry fused to	**VP16**	**NLS**
number of cells n	13	12
dissociation rate k (min^{-1})	0.07±0.04	0.08±0.03
m	1.53±0.20	1.47±0.22
half-life time $t_{1/2}$ (min)	4.9±0.8	4.8±0.6

Single dissociation curves were fitted to an exponential model with a time-dependent rate ($E = a \cdot e^{-k \cdot t_m} + c$) similar to a previously proposed model (Sing *et al.* 2014). The parameters correspond to the average of the single fit results and are listed with their respective standard deviations.

Taken together, BLInCR allows to accumulate effector proteins at genomic loci on the time scale of seconds after blue light illumination and those effectors dissociate on the time scale of minutes in the dark.

II.3.3 VP16 BLInCR activated the U2OS 2-6-3 transcription reporter system and possibly GSAT transcription at mouse PCH

Effector proteins such as the viral transactivator VP16 can be recruited to the vicinity of a minimal CMV promoter in the U2OS 2-6-3 reporter cell line, thereby activating transcription of the reporter. Active transcription can be tracked on the RNA level by transfecting the cells with fluorescently tagged MS2 coat proteins (MCP) that bind MS2 loop RNA or on the protein level by observing peroxisomal CFP signal (Janicki *et al.* 2004). To test if reporter transcription can efficiently be activated by VP16 BLInCR, the cells were transfected with an array marker (CFP-LacI), an array 'localizer' (CIBN-TetR), the PHR-YFP-tagged effector (VP16) as well as an RNA reader (mKate2-MCP) and illuminated

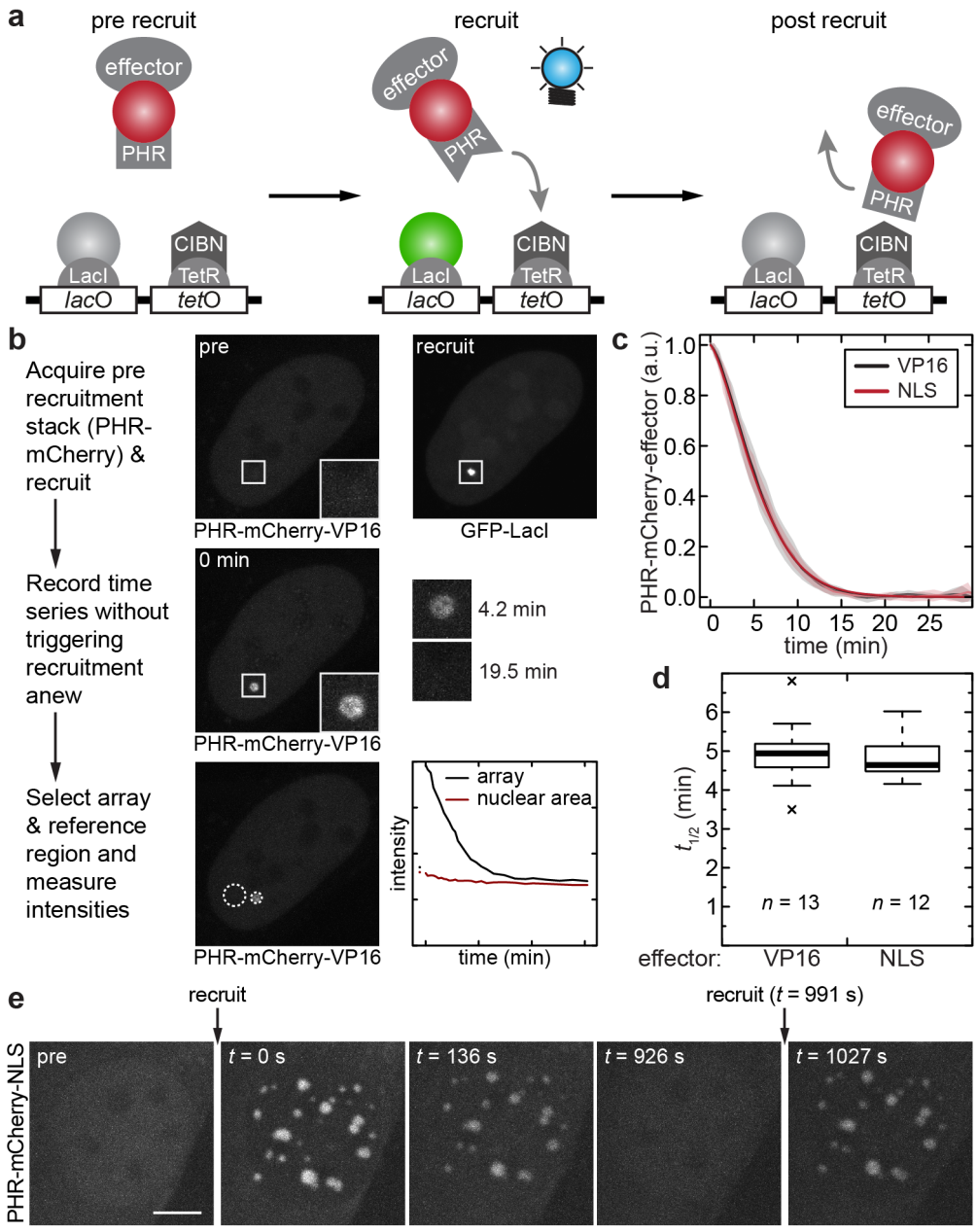

Figure 26: BLInCR dissociation kinetics. (**a**) Schematic representation of recruitment and dissociation. The array marker GFP-LacI was only imaged in the recruitment step and imaging the effectors did not lead to recruitment. (**b**) Workflow for BLInCR dissociation kinetics measurements. See text for details. (**c**) Measured dissociation kinetics for two effectors (VP16 and NLS). Averaged experimental data (transparent lines) with standard deviations (shaded areas) and averaged fits (solid lines) are shown, fit parameters are listed in **Table 5**. (**d**) Characteristic half-life for the two effectors. (**e**) Recruitment, dissociation and second recruitment of a mock effector (PHR-mCherry-NLS) to mouse PCH via CIBN-EYFP-mSUV4-20H2. BLInCR was activated by imaging YFP fluorescence in the localizer (not shown). Scale bar: 5 µm. Parts of the figure are adapted from (Rademacher *et al.* 2017).

with a blue LED over night (**Fig. 27a**). Fluorescently tagged MCP was enriched at the array and peroxisomal CFP could also be detected (**Fig. 27b**). In addition, accumulation of the largest subunit of pol II, RPB1, was visualized via a GFP-tagged construct (Sugaya *et al.* 2000) and could be detected at the array alongside the recruited activator (**Fig. 27c**). Transcription activation worked reliably with BLInCR as ∼90% of the cells showed MCP enrichment at the array and peroxisomal CFP could be detected in ∼85% of the cells, which was comparable to activation with FP-TetR-VP16 fusion constructs and more efficient compared to VP16 BLInCR to *lac*O (**Fig. 27c**). To adapt the system to mouse PCH in iMEFs, the tightly bound PCH component SUV39H1 was used as a 'localizer' to tether VP16 upon blue light illumination (**Fig. 27d**). VP16 BLInCR worked with this system (as shown in **Figure 24c** for NIH3T3 cells), but GSAT RNA cannot be detected using a live cell reader domain. Thus, transiently transfected cells were lysed after two hours of illumination with a blue LED and RNA levels were quantified using qRT-PCR. Preliminary results of two biological replicates revealed a small and insignificant increase of GSAT RNA compared to untransfected cells (**Fig. 27e**). However, the transfection efficiency in a comparable control experiment (same transfection protocol, constructs, cell line) where cells were fixed and imaged was only about 6%. Analyzing the bulk of transfected and untransfected cells might thus not be feasible under those non-optimized conditions. Nonetheless, VP16 BLInCR at an endogenous heterochromatic locus is technically possible and might overwrite the heterochromatic silencing and activate transcription of GSAT RNA. In the artificial reporter system of the U2OS 2-6-3 cell line, VP16 BLInCR to *tet*O reliably activated transcription that can be detected in single living cells.

II.3.4 Transcription output depended on the mode of activation and was heterogeneous in U2OS 2-6-3 cells

Reporter transcripts were strongly enriched in cells transfected with the relevant BLInCR constructs (CIBN-TetR, PHR-YFP-VP16 and mCherry-MCP) compared to mock-transfected (TetR-mRFP and YFP) or untransfected U2OS2-6-3 cells after over night exposure to blue light as determined by qRT-PCR (**Fig. 28a**). However, this bulk quantification does not offer valuable clues about the activation hetero- or homogeneity among single cells. Even though single cell microscopy can provide valuable insights on the heterogeneity of the response (**Fig. 28b**), exact quantification of MS2 RNA at the array across different cells and different constructs was not possible as the enrichment depends on the transfection levels of the effector constructs and mKate2-MCP. To account for those experimental differences, the average radial fluorescence intensity around the array was quantified, normalized to the mean fluorescence intensity of a nuclear reference area and averaged. The resulting radial profiles (**Fig. 28c**) of the recruited effector revealed that RNA enrichment at the array was always higher when recruiting to *tet*O compared to *lac*O

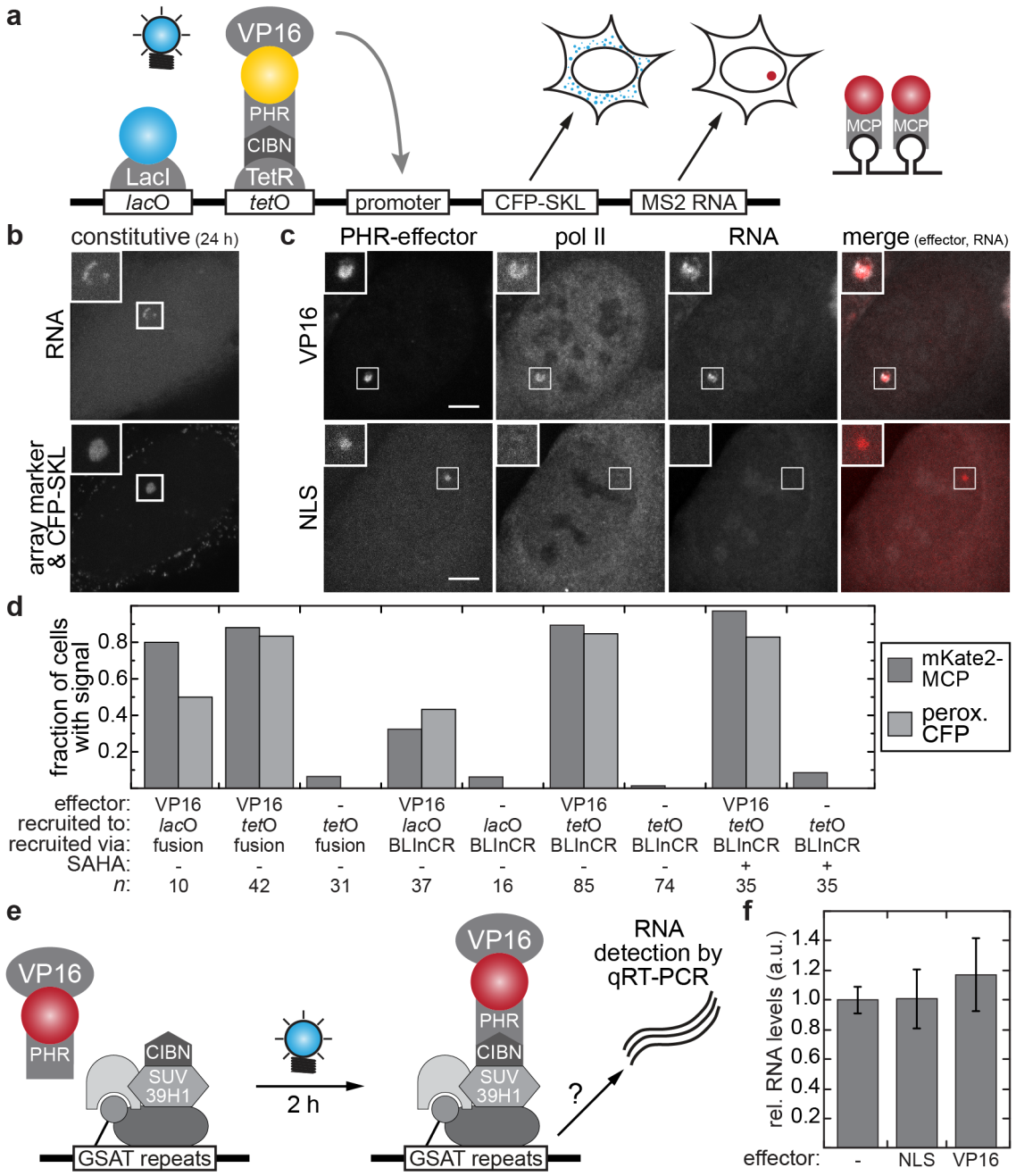

Figure 27: Transcription activation in U2OS2-6-3 cells and iMEFs. (**a**) Schematic representation of reporter transcription activation in U2OS 2-6-3 cells using BLInCR. (**b**) Reporter transcription was activated with the system depicted in (**a**) by illumination with a blue LED over night. RNA accumulation at the array (via mKate2-MCP) and peroxisomal CFP were readily detected in live cell CLSM images. CFP-LacI was used as an array marker. (**c**) CLSM images of pol II (GFP-RPB1) in cells activated as depicted in (**a**). Note that PHR constructs were labeled with iRFP713 and MCP with tagRFP to allow spectral separation of all constructs. (**d**) Fraction of cells that show reporter signal (RNA/peroxisomal CFP) across all experiments. (**e**) Schematic representation of VP16 BLInCR to activate GSAT transcription. CIBN-SUV39H1 was used as a 'localizer', but GSAT RNA cannot be detected by a fluorescently tagged reader domain. Instead, qRT-PCR was used. (**f**) Relative RNA levels after VP16 BLInCR as depicted in (**e**). Relative GSAT RNA levels after VP16 BLInCR determined by qRT-PCR. GSAT RNA levels were normalized using actin levels and untransfected controls. The mean ± sem sem of two biological replicates with three technical replicates each are shown. Parts of the figure were adapted and modified from (Rademacher *et al.* 2017).

(**Fig. 28c**, bottom, **Fig. 28b,d**), even though VP16 was recruited equally well. This effect was more pronounced for the BLInCR recruitment, which only activated <50% of the cells in the case of *lac*O (**Fig. 27d**). The integrated radial intensity from the array center to the full width at half maximum (FWHM) of the recruited effector was used as a measure for the accumulated RNA from radial profile plots (**Fig. 28d**). Some of the control cells that were not transfected with a VP16 construct also showed increased fluorescent signal at the array (**Fig. 27d**). However, this occurred rarely and quantifying the absolute fluorescence accumulation at the array $(I_{\text{array}} - I_{\text{ref}}) \cdot A_{\text{array}}$ in comparable cells that were recorded on the same day with the same laser intensity revealed that it was negligible compared to the MCP accumulation in the presence of VP16 (**Fig. 29a**, top). This was validated by RNA fluorescence *in situ* hybridization (FISH, **Fig. 29a,b**). Direct RNA visualization by FISH also showed that the heterogeneous transcription response to VP16 across different cells was not a transfection artifact of MCP, but actual cell-to-cell heterogeneity (**Fig. 29a**). In addition, the RNA level at the array (as detected by RNA FISH) was rather uncorrelated with the PHR-YFP-VP16 accumulation at the array (**Fig. 28**) suggesting that the array was saturated by VP16.

Lastly, the C-terminal domain (CTD) of pol II could be detected in cells activated by VP16 BLInCR over night, but not in non-activated cells transfected with a mock effector. In particular, the serine 5-phosphorylated form that is associated with transcription initiation (Stasevich *et al.* 2014) was strongly enriched (**Fig. 30**). Conversely, the elongation-associated (Stasevich *et al.* 2014) serine 2-phosphorylated CTD was only slightly enriched at the array (**Fig. 30**).

Characterization of the different recruitment and activation strategies showed that activation via recruitment to *lac*O was less efficient compared to *tet*O in U2OS 2-6-3 cells. This effect was more pronounced when using BLInCR compared to fusion constructs which might be due to interference with CFP-LacI which was always co-transfected in the BLInCR experiments (as an independent array marker). No additional array marker was used in the experiments with fusion constructs. Taken together, transcription of the reporter in the U2OS 2-6-3 cell line can be efficiently activated by VP16 BLInCR to tetO. As expected, pol II was enriched as shown by IF staining using this activation strategy.

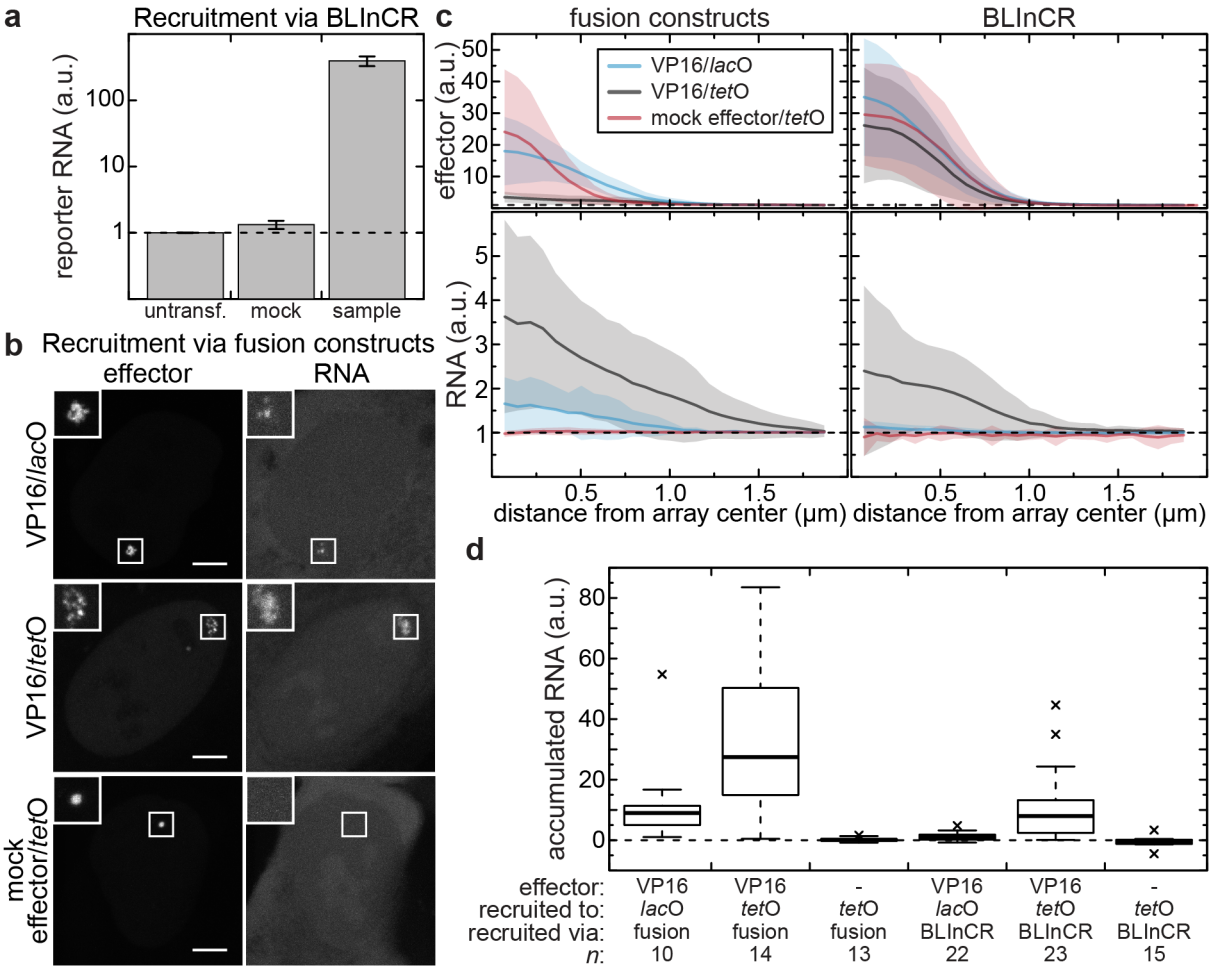

Figure 28: RNA quantification at the array after constitutive transcription activation. (**a**) qRT-PCR of untransfected and transfected U2OS 2-6-3 cells after constitutive activation over night. The mock transfected samples were transfected with TetR-mRFP and YFP and the sample was transfected with CIBN-TetR, PHR-YFP-VP16 and mCherry-MCP. Notably, only 10-15% of the cells showed red and yellow fluorescence (estimation by J. Trojanowski). Quantities were normalized using actin levels and the average value of untransfected cells. Error bars correspond to sem of two biological replicates (each with three technical replicates). (**b**) Live cell CLSM images of GFP-tagged fusion constructs (i. e. GFP-LacI-VP16, GFP-TetR-VP16, TetR-GFP; from top to bottom) and RNA (mKate2-MCP) at the array. Scale bars: 5 µm. (**c**) Average radial profiles of effectors (top) and RNA (bottom, mKate2-MCP reader) for recruitment via fusion constructs (left, as in (**b**)) or BLInCR (right, i. e. PHR-YFP-VP16/CIBN-LacI, PHR-YFP-VP16/CIBN-TetR, PHR-YFP/CIBN-TetR). The radial intensity $I(r)$ at distance r from the array center was normalized to the mean fluorescence intensity of a nuclear reference area I_{ref} (dashed line): $I_{\mathrm{norm}}(r) = I(r)/I_{\mathrm{ref}}$. The number of cells for each condition is listed in panel (**d**). (**d**) Accumulated RNA at the array calculated from the radial profiles in (**c**): $E_{\mathrm{sum}} = \sum_{r=0}^{r=FWHM} 2\pi r I_{\mathrm{norm}}$.

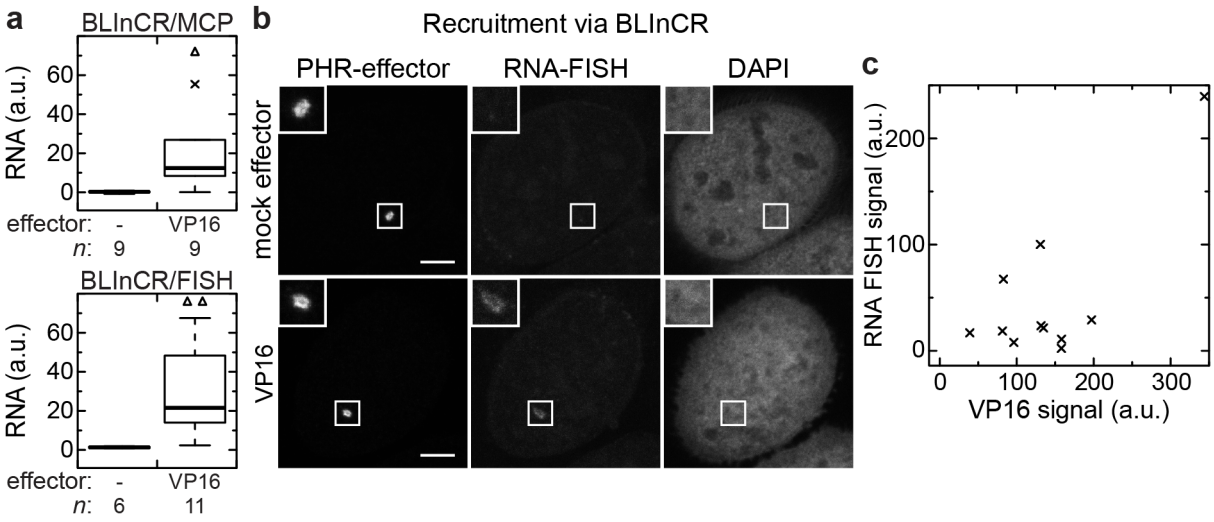

Figure 29: Different RNA visualization techniques were compatible with BLInCR and yielded comparable results. (**a**) RNA accumulation at the array after effector BLInCR over night. The RNA enrichment was calculated as $(I_{\text{array}} - I_{\text{ref}}) * A_{\text{array}}$ with I_{array} being the mean fluorescence intensity of a circular region around the array and A_{array} its area. Since all images for each experiment were recorded on the same day with the same laser intensities, this difference corresponds to the absolute mKate2-MCP or MS2 FISH probe accumulation at the array, respectively. The two experiments are not directly comparable with one another (different fluorescent probes). (**b**) Exemplary CLSM images of formaldehyde-fixed cells with MS2 RNA FISH using an Atto565-labeled probe after effector BLInCR. PHR-effector constructs were YFP-tagged and detected by anti-GFP IF (with an Alexa488-labeled secondary antibody) since YFP was destroyed during the FISH procedure. Scale bars: 5 µm. (**c**) Relationship between RNA FISH signal (from panel (**a**), bottom) and VP16 signal calculated analogously. Parts of the figure were adapted and modified from (Rademacher *et al.* 2017).

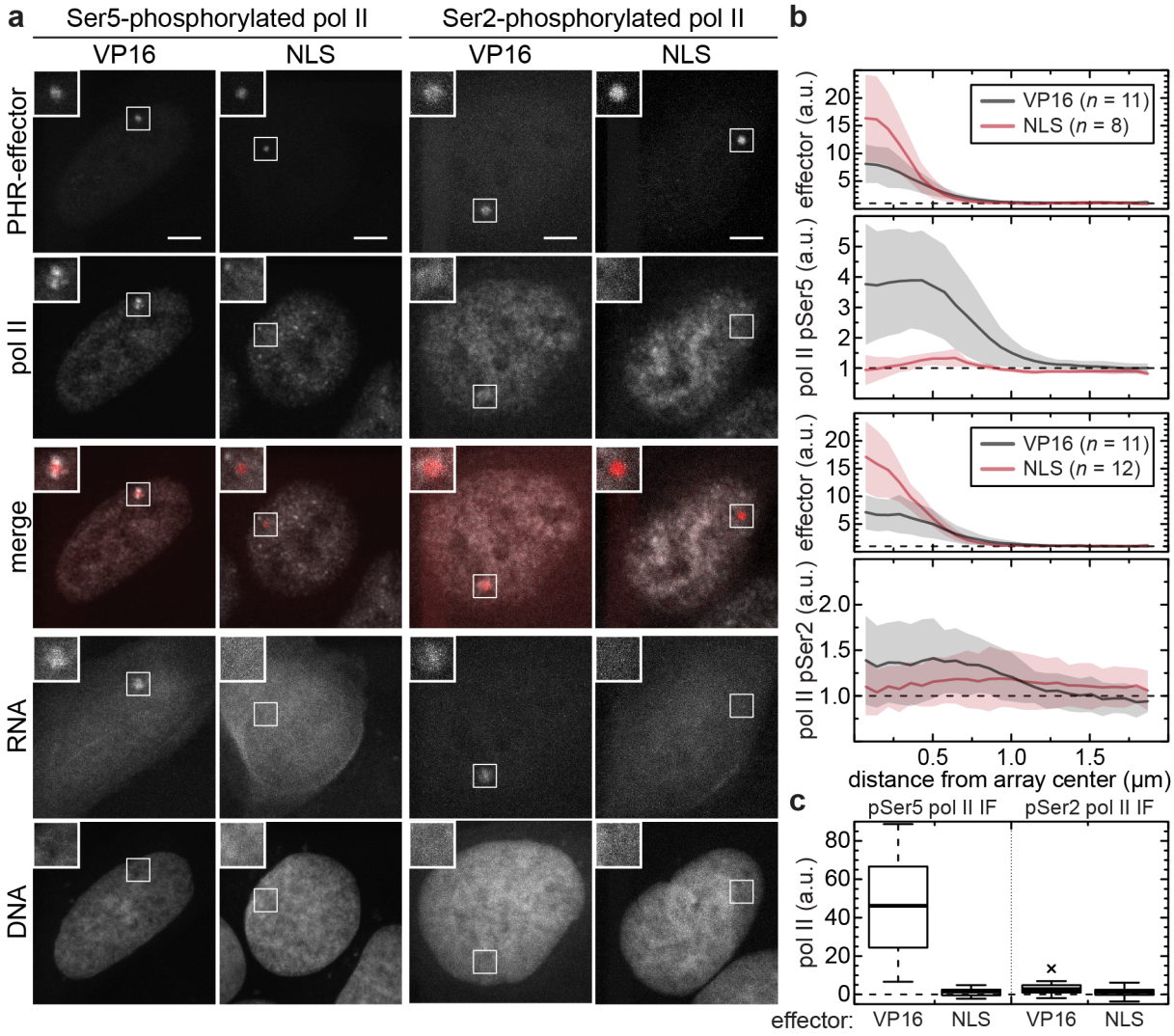

Figure 30: Immunofluorescence staining of DNA-directed RNA polymerase II after constitutive VP16 BLInCR. U2OS 2-6-3 cells were transfected with CIBN-TetR, PHR-YFP (mock) effectors (NLS or VP16) and mKate2-MCP and illuminated with a blue LED over night. (**a**) CLSM images of the PHR-effector constructs (YFP-tagged), IF of pol II phosphorylated at serine 5 or serine 2 (Alexa647-linked secondary antibody), merge of effector (red) and pol II (grey scale), RNA (mKate2-MCP) and DNA (DAPI) from top to bottom. Scale bars: 5 µm. (**b**) DAPI-normalized radial profiles of the effector and the modified pol II (mean and sd are displayed). (**c**) Accumulated pol II at the array. Calculations for (**b**&**c**) were done as for RNA described in **Figure 28c**&**d**.

II.3.5 Histone modifications and chromatin structure changed after long-term activation

Transcription activation in U2OS 2-6-3 cells has been reported to be accompanied by chromatin decondensation and histone acetylation (Rafalska-Metcalf *et al.* 2010), whereas non-activated cells displayed heterochromatic features such as binding of SUV39H1 and HP1 proteins as well as increased H3K9me3 (Janicki *et al.* 2004). U2OS 2-6-3 cells were transfected with the VP16 BLInCR system as depicted in **Fig. 27a**, constitutively activated and stained for acetylation and trimethylation of H3K9 by immunofluorescence to analyze the above-mentioned features of the silent and activated state. In activated cells, acetylation of H3K9 was rather evenly distributed across the array, but DAPI was slightly depleted (**Fig. 31a**, left) resulting in a small net enrichment seen in the DAPI-corrected profiles (**Fig. 31a**, right). In contrast, the negative control (no effector, PHR-YFP only) showed slight depletion of H3K9ac and homogeneous DAPI signal across the array. In both activated and control cells, a moderate local enrichment of H3K9me3 was observed in most cells (**Fig. 31b**), indicating that H3K9me3 and transcription were uncorrelated even though MS2 RNA production was not analyzed concomitantly.

Array decompaction was observed after long-term reporter transcription by expressing TetR-VP16 fusion constructs (**Fig. 28b**) and was also detected in the width of the radial array profiles (**Fig. 28c**, top left). However, assessing decompaction of the array upon activation by VP16 BLInCR is not directly feasible since PHR constructs oligomerize (Bugaj *et al.* 2013) possibly obscuring the actual array dimensions (**Fig. 28c**, top right). Therefore, CFP-LacI was co-transfected as an independent marker and the array size was determined in a threshold-based approach (see Materials & Methods for details). Activated arrays were very heterogeneous in size and RNA signal was detected at both small and large arrays (**Fig. 32a**). Recruitment of VP16-TetR fusion constructs led to pronounced decompaction, whereas activation via BLInCR enlarged the array to a slightly lesser extent (**Fig. 32b**). In summary, arrays tended to decompact upon constitutive activation, but decondensation of the locus was not a prerequisite for activation.

II.3.6 Visualizing epigenetic signatures in live cells with fluorescently tagged reader domains

Immunofluorescent staining of histone modifications (as in **Fig. 31**) is a standard technique to characterize the epigenetic signature of repetitive genomic loci, but it is inapt to trace changes in histone marks over time in living cells. In contrast, naturally occurring 'reader' proteins or protein domains can be fluorescently tagged and transfected into living cells. One such 'reader', namely murine bromodomain containing protein 2 (mBRD2), has been described to bind transcriptionally active arrays in the U2OS 2-6-3 cell line (Rafalska-Metcalf *et al.* 2010). To test whether it is compatible with VP16

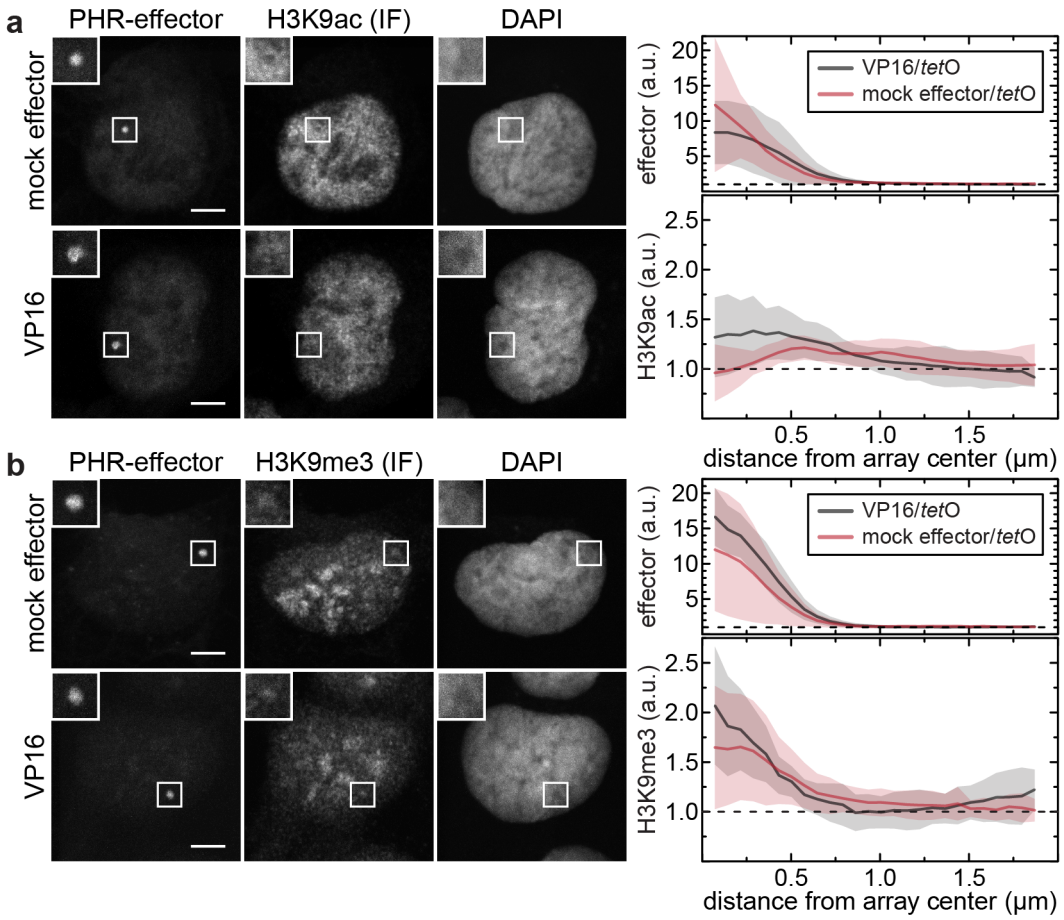

Figure 31: Immunofluorescence staining of histone modifications after constitutive VP16 BLInCR to *tet*O. U2OS 2-6-3 cells were transfected with CIBN-TetR and PHR-YFP-VP16 or PHR-YFP only. CLSM images of YFP fluorescence from the recruited effector (left), immunofluorescence staining of histone modifications (center, Alexa568-linked secondary antibodies) and DAPI staining (right) of fixed cells after blue light illumination over night are displayed. (**a**) Acetylation of H3K9. (**b**) Trimethylation of H3K9. The profiles on the right were normalized to a nuclear reference area as described above and the respective DAPI profile. For each condition, $n = 10$ cells were analyzed except for the negative control H3K9me3 staining without VP16 ($n = 11$). Scale bars: 5 µm. Parts of the figure were adapted and modified from (Rademacher *et al.* 2017).

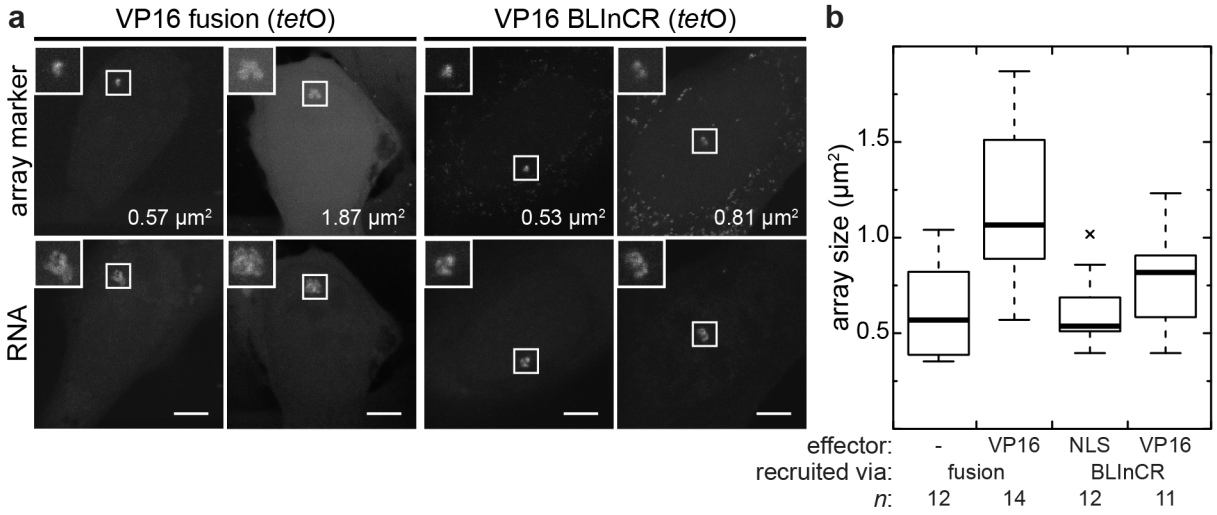

Figure 32: Array size after constitutive transcription activation. (**a**) CLSM images of the independent array marker (CFP-LacI) and RNA (mKate2-MCP). Cells were either activated with a VP16-TetR fusion construct or VP16 BLInCR and were heterogeneous in size. Scale bars: 5 µm. (**b**) Array size quantified from CLSM images by a threshold-based approach. The figure was adapted from (Rademacher *et al.* 2017).

BLInCR, it was cotransfected with the BLInCR components depicted in **Figure 27**. A systematic analysis of recruitment to either *lac*O or *tet*O with fusion constructs or VP16 BLInCR revealed that recruiting VP16 fusion constructs to *lac*O resulted in more mBRD2 enrichment compared to *tet*O (**Fig. 33**). This indicates that mBRD2 recruitment might be independent of transcription activation since recruitment to *tet*O generally resulted in higher transcript levels (**Fig. 28**), bbut lower mBRD2 levels. BLInCR of VP16 or a mock effector to *tet*O showed no differential mBRD2 binding at the array, suggesting either that mBRD2 did not efficiently bind acetylated histones in the given chromatin context or that histones at the array were not acetylated despite transcription activation. Another bromodomain containing protein, mBRD4, has been described to be co-recruited along VP16 (Rafalska-Metcalf *et al.* 2010) and showed strong enrichment at the array when cells were transfected with VP16 fusion constructs. As for mBRD2, activator recruitment to *tet*O resulted in lower mBRD4 enrichment compared to *lac*O (**Fig. 33**).

The chromodomain of mouse HP1β (mCD) and the human ADD domain of the ATRX protein bind H3K9me3 (Bannister *et al.* 2001, Dhayalan *et al.* 2011) and might also be suitable to be used as reader domains in living cells as reported for the mCD (Muller-Ott *et al.* 2014). Here, those two domains were tested as live cell readers to assess the H3K9me3 status of the array under activated and silenced conditions. When no activator is recruited, the reporter locus in the U2OS 2-6-3 cell line has been described to be trimethylated at H3K9 and enriched with heterochromatin-associated proteins such as the histone methyl transferases SUV39H1 and G9A as well as HP1 (Janicki *et al.* 2004). These findings were confirmed by the immunostainings in **Figure 31b** above.

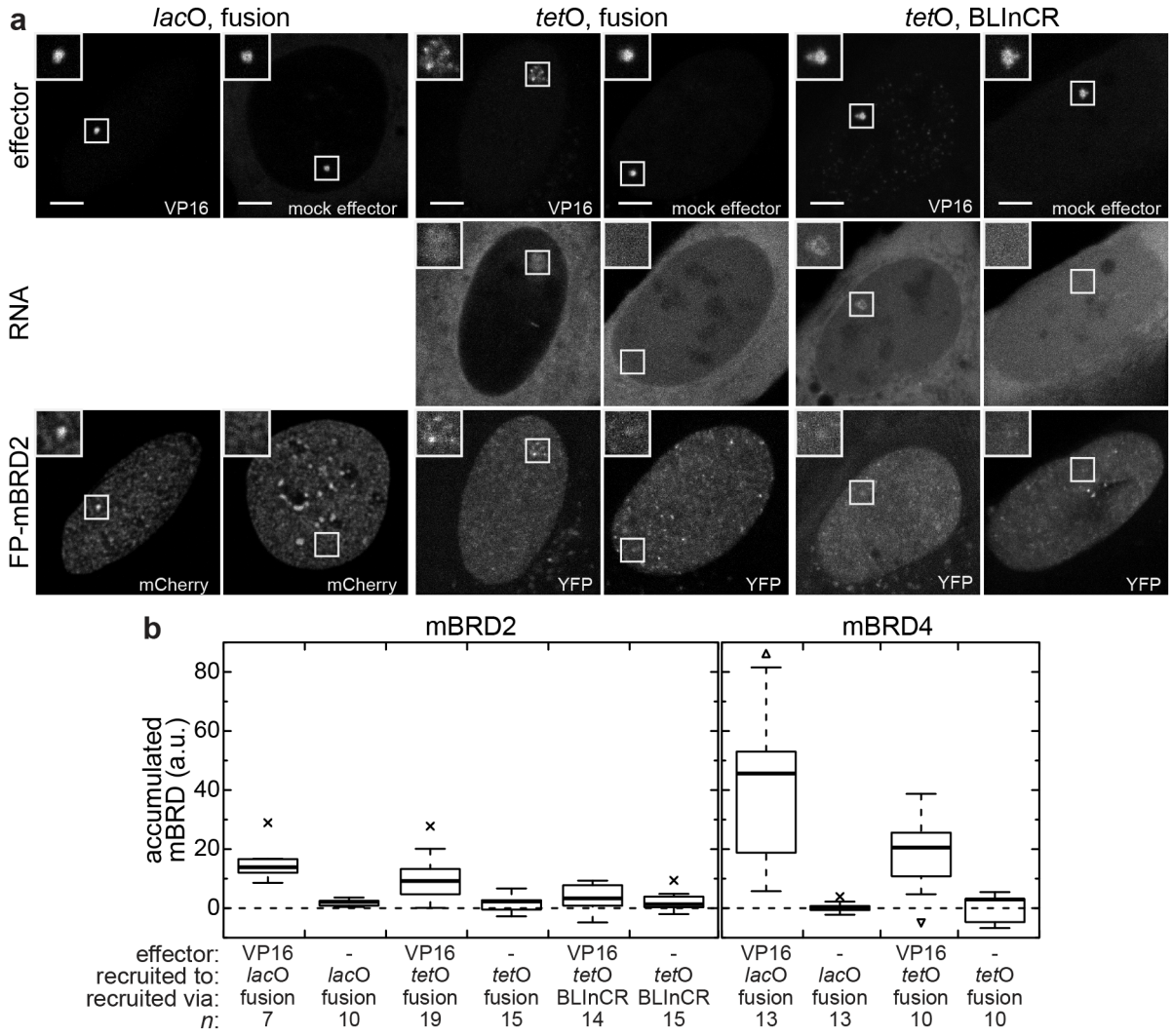

Figure 33: Bromodomain-containing protein accumulation at the U2OS 2-6-3 reporter array for different activation modes. (**a**) Live cell CLSM images of recruited effector (top), RNA reporter (mCherry-MCP, middle) and fluorescently tagged mBRD2 (bottom). The fusion effectors were tagged with tagBFP, whereas the BLInCR effectors were tagged with iRFP713. Scale bars: 5 µm. (**b**) Quantification of accumulated mBRDs at the array. The quantification of accumulated mBRD at the array was based on the radial profiles as in **Figure 28d**.

The H3K9me3 reader domains were thus tested under the silenced conditions by marking either the *tet*O or the *lac*O repeats with fluorescently tagged TetR or LacI constructs, respectively. The mCD was enriched at the array in most cells when marking the array at the *lac*O sites, whereas enrichment of the hADD domain was observed less frequently and at lower levels. In addition, reader enrichment was always lower when the array marker was recruited to *tet*O compared to *lac*O (**Fig. 34a**). Notably, this effect was observed both with the mBRD proteins and the H3K9me3 reader domains, suggesting that it might be related to the general chromatin landscape at the array. Transcription activation led to a slight reduction of average H3K9me3 levels at the array (**Fig. 34b**), but nonetheless cells with mCD enrichment at the array were observed even though they were activated over night (**Fig. 34b**, top). This is consistent with the antibody staining of H3K9me3 after constitutive activation shown in **Figure 31**.

In summary, the localization of live cell readers of histone modifications did not only depend on the activation state of the reporter array in the U2OS 2-6-3 cell line, but also on the mode of activation (BLInCR vs. fusion constructs) and the site to which the (mock) effectors were recruited (*lac*O vs. *tet*O). VP16 BLInCR did not lead to differential enrichment of mBRD2 at the array, but did lead to a slightly lower average mCD enrichment compared to the mock control. Interestingly, activated cells with some mCD accumulation at the array were observed, confirming the finding that H3K9me3 alone cannot repress transcription as shown for PCH in NIH3T3 cells (**Fig. 18a**) and some genes in wt iMEFs (**Fig. 21c**).

II.3.7 Transcription activation via VP16 BLInCR was rapid and accelerated in cells with hyperacetylated histones

The VP16 BLInCR setup in the U2OS 2-6-3 reporter cell line depicted in **Fig. 27a** can reliably activate transcription and detect RNA in cells. It is thus suited to follow reporter activation by time-lapse live cell imaging. To analyze the kinetics of transcription activation in detail, VP16 was constantly re-recruited by simultaneous imaging the array marker (CFP-LacI), the effector (PHR-YFP-VP16) and the RNA reader (mKate2-MCP) every 2-5 min. In addition, cells were cultured either with or without the histone deacetylase inhibitor SAHA prior to live cell imaging to assess the role of hyperacetylation in transcription initiation. Treatment with 2 µM SAHA for 24 h led to an increase in histone acetylation at H3K9, H3K27 and H4 as shown by western blot analysis (**Fig. 35a**). Concordantly, constitutively activated cells (with/without SAHA treatment) stained for H3K9ac by IF revealed that the accumulated fluorescence signal in the entire nucleus increased in SAHA-treated cells compared to untreated cells. Direct comparison of fluorescence intensities was feasible for this analysis, because cells were cultered in removable 4-well chambers and subsequently stained on the same slide and imaged with the same parameters. Increased acetylation was also seen at the array (**Fig. 35b**). For this analy-

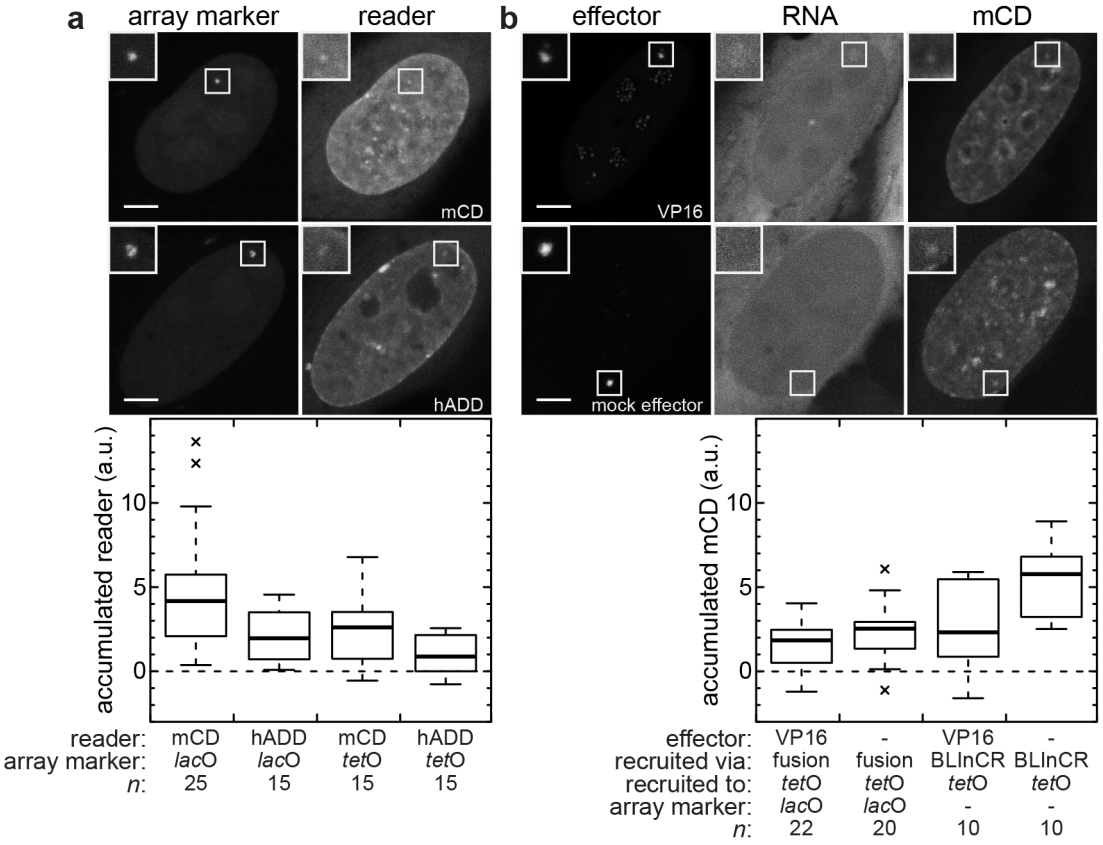

Figure 34: Live cell H3K9me3 reader domains (**a**) The chromodomain of mouse HP1β (mCD) and the human ATRX-Dnmt3-Dnmt3L (hADD) domain of the ATRX protein were enriched at the non-activated U2OS 2-6-3 reporter array albeit at different amounts. The quantification below the CLSM images was based on the radial profiles as described for **Figure 28d**. (**b**) Activation of the array via VP16 only led to a slight reduction of mCD accumulation at the array compared to the controls without effector. Note that CFP-MCP did not work well as an RNA reader (top row, center), but all cells showed either a slight enrichment and/or CFP-SKL confirming activation. Scale bars: 5 µm.

sis, the integrated intensity from the array center up to the FWHM was quantified (as in **Fig. 28d**). Interestingly, array acetylation increased stronger in activated cells (VP16) compared to control cells (NLS) suggesting that pre-existing hyperacetylation might further enhance activation-associated acetylation that was only moderate in the absence of SAHA (**Fig. 31a**).

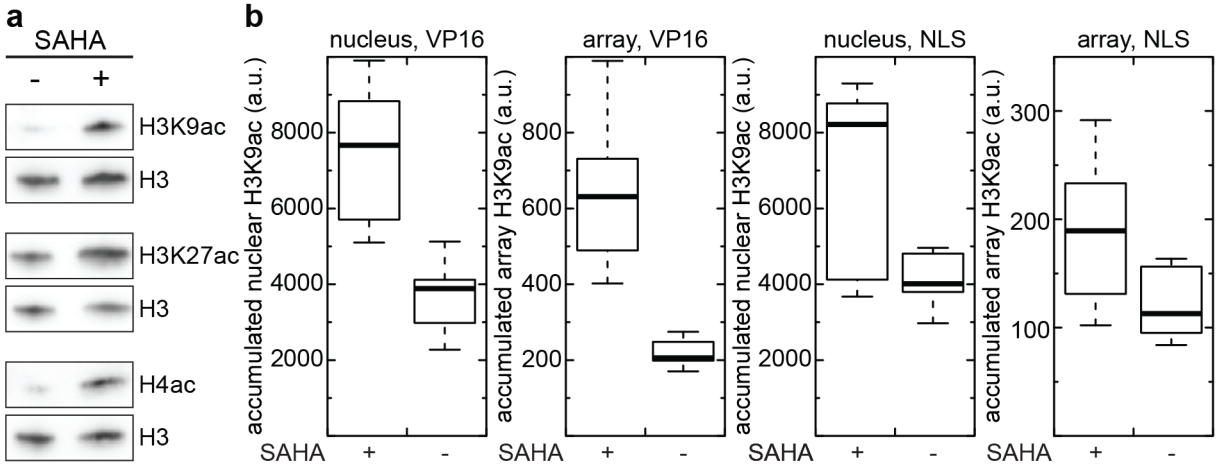

Figure 35: Treatment with the histone deacetylase inhibitor SAHA led to histone hyperacetylation in U2OS2-6-3 cells. (**a**) Western blot analysis of acetylation of H3K9, H3K27 and H4 in the presence and absence of SAHA. After detecting the acetylation signals, antibodies were stripped from the membrane and H3 was detected as a loading control. (**b**) Quantification of fluorescence signal after IF staining for H3K9ac in transfected cells. Cells were grown and stained on the same slide (with a removable 4-well chamber) and imaging was done on the same day using the same imaging parameters, thereby allowing for direct comparison of fluorescence intensities. For the nuclei quantification, the mean nucleus signal was multiplied with the area of the nucleus. For the array quantification, the signal was integrated from the array center to the FWHM (as determined in the effector channel). The array signal was normalized to the DAPI ratio (array/whole cell) to correct for changes in chromatin structure – i. e. decondensation of the array and associated DAPI decrease observed in some of the activated cells. 12/6 treated and 12/8 untreated cells were analyzed (VP16/NLS). The figure was adapted and modified from (Rademacher *et al.* 2017).

In some cells, RNA was already detected in the second frame, i. e. after 2 min (**Fig. 36a**). Activation generally proceeded in two phases: a rapid first activation leading to a plateau in RNA accumulation between 10-20 min after initial VP16 BLInCR followed by a second increase of transcript accumulation after 20-40 minutes (**Fig. 36b**). Hyperacetylation in cells pre-treated with SAHA resulted in a more pronounced first activation phase and a more rapid response to the activator in the sense that RNA was detectable after less than five minutes in all cells compared to the broad distribution of response times in untreated cells (**Fig. 36c**). In addition, the produced RNA was not distributed homogeneously across the array (**Fig. 36a**, insets) indicating that some of the ∼200 promoters might be poised for transcription and activated more rapidly than others.

During the first 50 min of activation, the array structure did not change on the scale that is accessible by light microscopy (**Fig. 36a**, middle, **Fig. 36b**, bottom). The average array decondensation observed after constitutive activation over night (**Fig. 32b**) was thus rather a downstream effect of transcription activation. The biphasic activation kinetics fitted a positive feedback model with a fraction of independent transcription units. A sequential activation model with n steps and identical transition rates between those steps also yielded a good, but slightly worse fit (**Fig. 36b**). Importantly, simulation of different recruitment speeds showed that the fast BLInCR recruitment and detection of RNA were necessary to resolve the biphasic activation kinetics (Rademacher *et al.* 2017).

Combining VP16 BLInCR and live cell RNA detection in the U2OS 2-6-3 reporter system allowed to monitor the onset of transcription in single living cells at high spatial and temporal resolution. The following results were obtained here from the application of this system: i) the transcription activation kinetics of a gene cluster proved to be biphasic; ii) the spatial organization of the array as accessible by confocal microscopy did not change during the first 50 min of activation; iii) RNA was produced heterogeneously across the array; iv) the transcriptional response to a strong activator was heterogeneous across different cells and v) treatment with the histone deacetylase inhibitor SAHA led to faster response times and a more pronounced first activation phase.

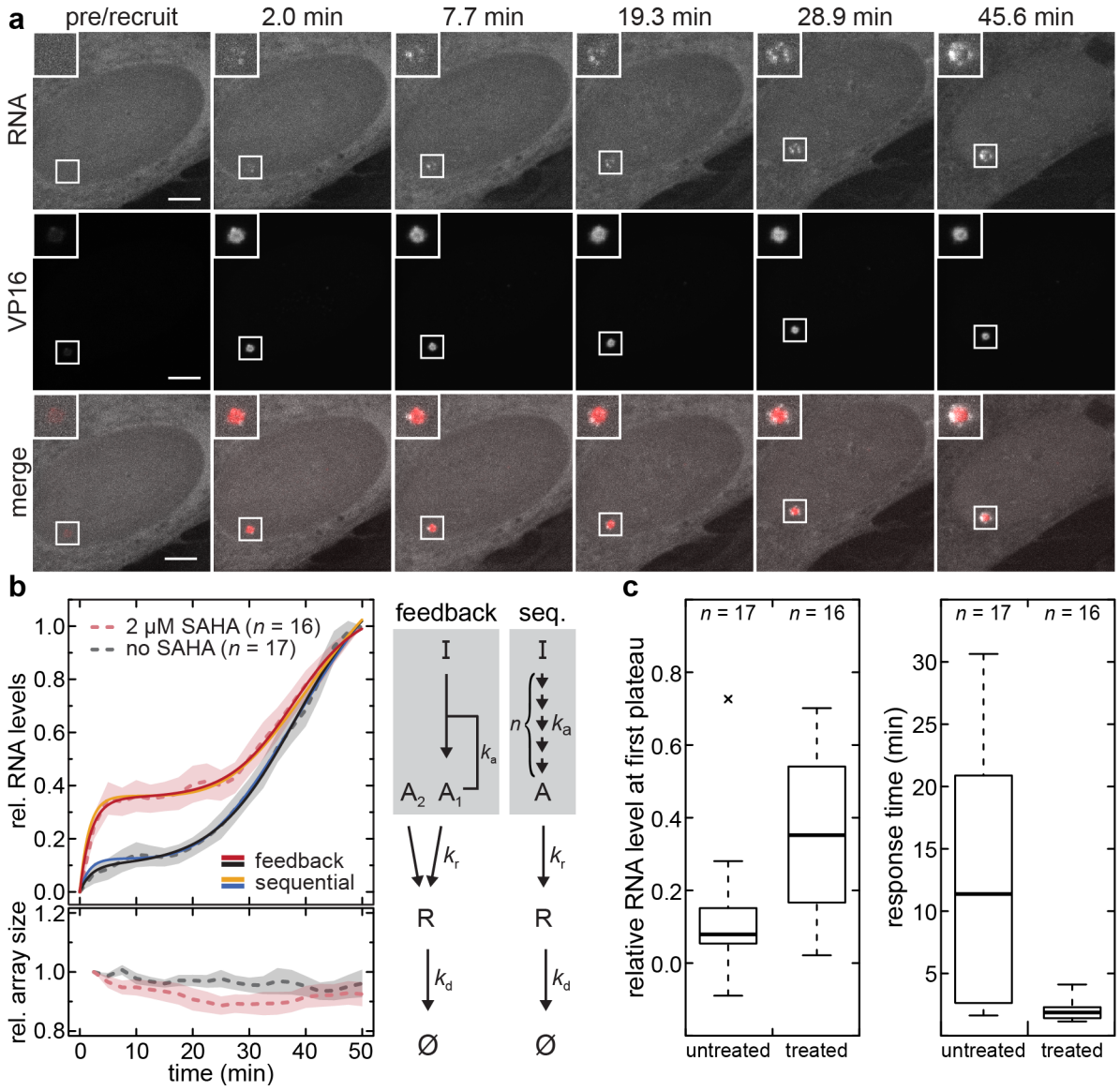

Figure 36: Transcription activation kinetics after VP16 BLInCR. (**a**) CLSM images of the RNA reader (mKate2-MCP, top) and the VP16 effector (PHR-YFP-VP16, middle) as well as a merge (bottom). Scale bars: 5 µm. (**b**) Relative RNA levels during the first 50 min of activation (top, corrected for fluorophore bleaching and normalized to the levels at t = 50 min). Bottom: relative array size compared to the second time point (i. e. when VP16 was fully recruited) during the activation. The RNA levels fitted a model including positive (black/red) or (slightly worse) a sequential activation model (yellow/blue). Modeling was done by F. Erdel. (**c**) Response time and relative RNA levels at the first plateau (i. e. at the point of inflection: 8.7 min for untreated and 11.8 min for treated cells) for the cells analyzed in (**b**). The figure was adapted and modified from (Rademacher *et al.* 2017).

II.3.8 The persistence of the activated state could be probed using BLInCR

BLInCR is a promising tool to study the reversibility of biological processes by taking advantage of the dissociation kinetics in the dark (**Fig. 26**). To test this system for the reversibility of transcription activation, the light trigger that targeted VP16 to the reporter was removed after an initial recruitment phase. If the transcriptionally active state were memorized even in the absence of the trigger (i. e. VP16), increased total reporter RNA levels (**Fig. 37a**, top) and detectable array-associated RNA would be expected. In the alternative case, transcriptional activity would not persist leading to a decrease in total reporter RNA levels (**Fig. 37b**, bottom) and dissociation of fluorescently tagged MCP from the array. Both cases were detected in live cell experiments. In some cells, PHR-iRFP713-tagged VP16 readily accumulated, activated transcription at the array after minutes and dissociated in the dark, causing transcription to decrease (**Fig. 37b**&**c**, cell 1). When recruiting VP16 anew, transcription increased again, possibly even faster than after the first BLInCR pulse. However, other cells were strongly activated with a single VP16 BLInCR pulse (52 s), which resulted in a stably activated state that was independent of light-induced VP16 tethering and irreversible on the time-scale considered here (**Fig. 37c**, cell 2). Notably, VP16 did not completely dissociate from the array in this case, which could be due to interactions with other proteins of the transcription machinery or with RNA. A bulk experiment according to the scheme in **Fig. 37a** was performed to assess which of those scenarios is more prevalent. Cells were transfected with CIBN-TetR and PHR-YFP-VP16 and illuminated for the indicated amount of time (t_{light}). Reporter RNA levels were measured by qRT-PCR at different time points after turning the light off (t_{dark}) and displayed a decrease over time suggesting that the fully activated state does not persist.

All in all, using BLInCR for transcription activation generated valuable insights on the early steps of transcription activation and provides a tool to assess the memory of the activated state. BLInCR was shown to be compatible with live cell detection of RNA as well as qRT-PCR. In combination with epigenetically modifying drugs and live cell monitoring of epigenetic readers and writers, BLInCR can be valuable in dissecting mechanisms of epigenetic regulation.

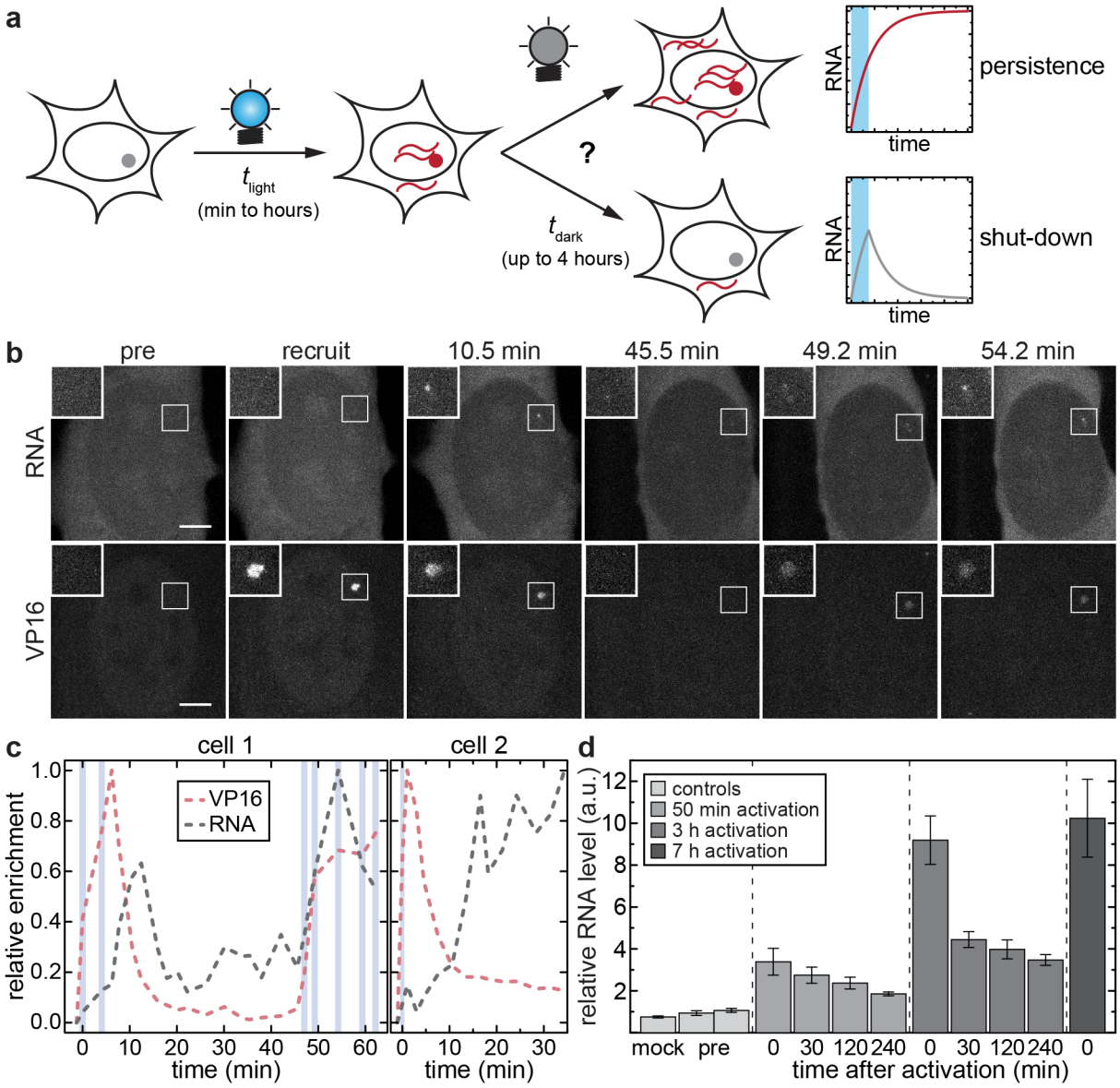

Figure 37: Reversible transcription activation by VP16 BLInCR and live cell imaging of RNA production. (**a**) Schematic representation of experimentally testable scenarios addressing the persistence of the transcriptionally active state. If the active state is memorized, total reporter RNA should keep on increasing after the activator is lost from the array (top). Alternatively, total reporter RNA levels would decrease if the activated state does not persist (bottom). (**b**) CLSM images of the RNA reader (tagRFP-MCP, top) and the effector (PHR-iRFP713-VP16, bottom). The activation pattern is depicted in panel (**c**). Scale bars: 5 µm. (**c**) Quantification of VP16 and RNA accumulation at the array. Enrichment values were corrected for bleaching and normalized to the measured maximum. The shaded blue areas represent the recruitment light pulses. Cell 1 is shown in panel (**b**). (**d**) Total reporter RNA levels after activation for 50 min, 3 h or 7 h (t_{light} in (**a**)) determined by qRT-PCR. The time after activation corresponds to t_{dark} in (**a**). Actin-normalized quantities are depicted relative to the average level before illumination ("pre"). The mock control corresponds to untransfected cells. The qRT-PCR was performed by S. Schumacher. Parts of the figure were adapted and modified from (Rademacher *et al.* 2017).

III Discussion

III.1 PCH specificity and memory is governed by chromatin-bound 'bookmarking' factors and possibly DNA methylation rather than by PAX transcription factors

Although PCH in MEFs is a thoroughly studied model system for constitutive heterochromatin, the question how PCH-specific stably bound complexes find their sites of action after DNA replication and cell division remains unresolved. The 234 bp major satellite consensus repeat sequence harbors a sequence somewhat similar to the PAX3 transcription factor binding site. It has therefore been suggested that PAX3 is the DNA sequence-specific factor that targets PCH and is necessary for efficient repression of GSAT transcription (Bulut-Karslioglu *et al.* 2012). PAX3-GFP showed a predominantly homogeneous distribution across the nucleus and was excluded from nucleoli (**Fig. 10a**,**b**). The heterochromatic localization of PAX3-GFP in one cell was probably a transfection artifact as this has only been observed in transfected cells and not when staining endogenous PAX3 as described elsewhere (Corry *et al.* 2008). The observed inability of PAX3 to tightly bind major satellite repeats in living cells might be explained by the mismatches of the apparent binding site within the major satellite consensus sequence compared to the Paired domain Recognition Site 1 (PRS-1) described for mouse PAX3 (Chalepakis *et al.* 1994) (**Fig. 38**).

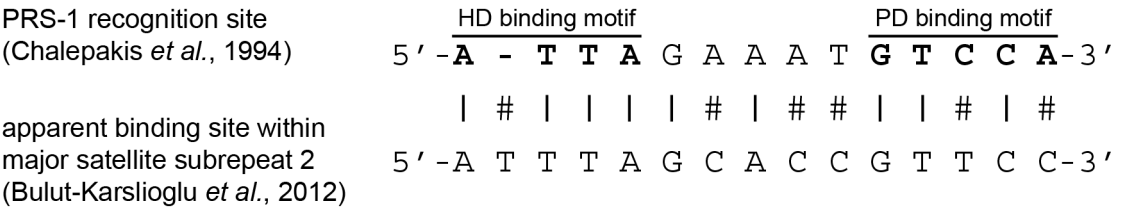

Figure 38: Alignment of published PAX3 binding motif and apparent Pax3 binding site within major satellite repeats. Mutations in the upstream paired type homeodomain (HD) or the paired domain (PD) binding motifs have been found to be sufficient to abolish PAX3 binding (Chalepakis *et al.* 1994), whereas the spacer between the two domains can vary in length. The absence of the correct HD or PD domain in the major satellite consensus sequence could explain the inability of PAX3 to consistently localize to PCH.

In addition, PAX3-GFP showed high mobility both at PCH and in euchromatic regions (**Fig. 10c**, **Table 1**), which is also consistent with previously published FRAP data (Corry *et al.* 2008). Notably, the apparently longer time until complete recovery observed by Corry and colleagues is likely due to the difference in bleach geometry (strip vs. circle FRAP). Taken together, PAX3 cannot be considered a component of the stably bound SUV39H-centered PCH-specific complex (Muller-Ott *et al.* 2014) since it is rather

depleted from PCH when normalized to DNA density (**Fig. 10b**) and not immobilized. However, PAX3 might contribute to induce GSAT transcription during some cell cycle phase (e. g. during S-phase), thereby recruiting PCH factors such as SUMOylated HP1 or SUV39H2 for *de novo* PCH establishment as suggested elsewhere (Maison *et al.* 2011, Velazquez Camacho *et al.* 2017).

An (epigenetic) factor upstream of SUV39H might also be responsible for targeting SUV39H/HP1 specifically to PCH. The methyl-CpG-binding protein MECP2 as well as 5-methyl-cytosine (5mC) remains enriched at DAPI-dense foci in the absence of SUV39H KMTs in iMEFs (Muller-Ott *et al.* 2014). Thus, the DNA methylation pathway harbors intriguing candidates for such a factor. Indeed, MECP2 has been shown to direct KMT activity specifically to H3K9 (Fuks *et al.* 2003b). Conversely, DNA methylation has been reported to strongly depend on H3K9me3 in *Neurospora crassa* and *Arabidopsis thaliana*, where loss of the respective KMT resembles the loss of DNA methyltransferase phenotype (Jackson *et al.* 2002, Tamaru & Selker 2001). Similarly, DNA methylation was reduced upon loss of SUV39H in mESCs (Lehnertz *et al.* 2003). In iMEFs, DNA methylation appeared to be independent of SUV39H, since MECP2 localization and mobility was unchanged in SUV39H dn compared to wt iMEFs (**Fig. 11**). MECP2 could therefore recruit SUV39H to PCH. This notion is consistent with the observation that reintroducing SUV39H dn iMEFs with SUV39H-GFP results in PCH enrichment of the GFP construct (Muller-Ott *et al.* 2014). Krouwels and colleagues also showed that SUV39H1 reintroduction into SUV39H dn iMEFs resulted in a rescue of the PCH state in the sense that H3K9me3 and HP1β enrichment were reestablished (Krouwels *et al.* 2005). The DNA methylation pathway thus appears to be upstream of the H3K9me3/SUV39H pathway at iMEF PCH rather than being part of a self-reinforcing cycle of repressed epigenetic states as proposed elsewhere (Fuks 2005). However, this raises the question how MECP2 specificity for PCH comes about.

MECP2 has been found to bind cooperatively to methylated DNA (Ghosh *et al.* 2010) and to have a slight preference for the non-methylated major satellite DNA sequence compared to other non-methylated competitor DNA sequences (Lewis *et al.* 1992). Still, the moderate 5mC enrichment at PCH compared to euchromatin ($\sim$7-fold) alone cannot explain the strong enrichment of immobilized MECP2 at PCH ($\sim$80-fold) (Muller-Ott *et al.* 2014). The strongly enriched immobilized MECP2 could be sufficient for SUV39H immobilization ($\sim$50/15-fold for SUV39H1/2, respectively). In addition, 2 out of 16 mappable interspersed major satellite repeats were reported to be active as judged by the presence of H3K36me3 and did not show SUV39H/HP1/H3K9me3 enrichment (Muller-Ott *et al.* 2014). Interestingly, those two repeats displayed a comparably high 5mC level as the silenced repeats (Muller-Ott *et al.* 2014). Thus, another factor than 5mC must contribute to MECP2 and – downstream – to SUV39H immobility. This hypothesis is indirectly supported by the finding that SUV39H1 mobility increases only slightly upon

DNA demethylation after treatment with 5-aza-C (Krouwels *et al.* 2005). Even though 5-aza-C treatment leads to a pronounced reduction of 5mC and MECP2 at PCH, a fraction of MECP2 remained rather immobilized (Ghosh *et al.* 2010) which might be indicative of tethering through another factor or binding to unmethylated chromatin as shown for human MECP2 *in vitro* (Georgel *et al.* 2003). To further explore the linkages between the 5mC/MECP2 and the SUV39H/HP1/H3K9me3 pathway, it would be interesting to explore if reintroduction of SUV39H into 5-aza-C-treated SUV39H dn iMEFs would also rescue the PCH state.

Finally, the possibility that SUV39H1 remains chromatin-bound throughout all cell cycle stages thereby acting as a 'bookmark' to denote major satellite repeats as sites to reestablish PCH after DNA replication and cell division was explored. These experiments revealed that a fraction of SUV39H1 remained indeed chromatin-bound at all times (**Fig. 12**). SUV39H1 was also detected at nascent chromatin right after DNA replication along with the DNA methyltransferase DNMT1 and the HP1 proteins (Alabert *et al.* 2014), which is critical for a 'bookmarking' factor. Notably, sparsely distributed tightly bound PCH complexes (**Fig. 39a**) are sufficient to (re)establish PCH domains (Muller-Ott *et al.* 2014). Thus, the high affinity complexes might remain chromatin-bound at major satellite repeats throughout the cell cycle and reestablish PCH after cell division, whereas the fraction of loosely bound components dissociates (**Fig. 39b**).

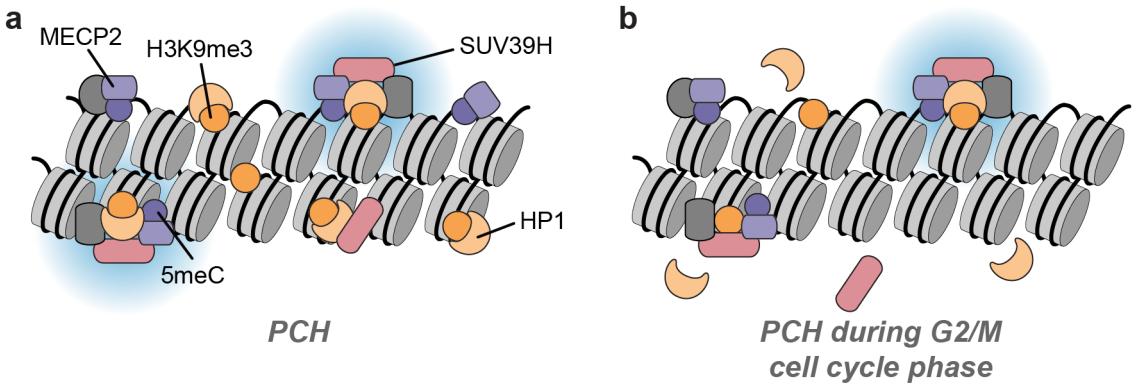

Figure 39: PCH-specific, stably chromatin-bound complexes centered around SUV39H/HP1/H3K9me3 and possibly 5mC/MECP2 remain chromatin-bound throughout the cell cycle. (**a**) Stably bound PCH complexes (blue background) alongside more transiently associated H3K9me3- and 5mC-binding proteins as observed during G1 phase. The gray non-histone protein next to MECP2 and 5mC could be an additional factor that stabilizes MECP2 at PCH. (**b**) PCH during G2 phase and mitosis. Rather loosely bound proteins dissociate and get diluted during cell division, whereas some stably bound PCH complexes remain chromatin-bound. Figure adapted and modified from (Muller-Ott *et al.* 2014).

III.2 Loss of SUV39H2 influenced global and local H3K9me3 levels

The current view is that the two isoforms of the SUV39H KMT play redundant roles in the establishment of H3K9me3 at least during embryonic development (O'Carroll *et al.* 2000, Peters *et al.* 2001). However, SUV39H2 had a greater influence on global H3K9me3 levels (western blot analysis, **Fig.17a**) and PCH H3K9me3 (microscopy and ChIP-seq, **Fig.17b-d**). The SUV39H2 sn cell line E480-2 showed intact global H3K9me3, but reduced PCH H3K9me3 (**Supplementary Fig. S1**), indicating that the effect on global H3K9me3 levels might have been compensated by other KMTs. Some SUV39H2-deficient cells displayed wt levels of H3K9me3 and HP1 at chromocenters (e. g. the upper right cell depicted in **Fig.17b**, row 3). Thus, SUV39H1 can in principle maintain intact PCH in the absence of SUV39H2, but might not always be targeted to pericentromers leading to a loss of H3K9me3 at PCH in some cells (e. g. the lower left cell depicted in **Fig.17b**, row 3). This is consistent with the previous finding that transfected SUV39H1 can reestablish PCH H3K9me3 in SUV39H dn iMEFs (Krouwels *et al.* 2005). Furthermore, a contribution of SUV39H1/2 to targeting other PCH components to pericentromeres has recently been proposed (Johnson *et al.* 2017, Velazquez Camacho *et al.* 2017). Taken together, both SUV39H1 and SUV39H2 seem to be able to set H3K9me3 in living cells. SUV39H2, however, might have a more prominent role in targeting pericentromers, thereby guiding other PCH components. A detailed analysis of reintroduced SUV39H constructs into SUV39H dn cells and their efficiency in correctly localizing to PCH and reestablishing H3K9me3 might be a starting point to uncover the cause of the heterogeneity observed even within one cell line (as shown in **Fig.17b**, row 3). All but one of the 23 mappable, interspersed GSATs analyzed here displayed high levels of H3K9me3 (>3-fold above genomic average) in wt iMEFs and even higher levels in mESCs. All normalized GSAT H3K9me3 levels were in the range of 2.5- to 4.5-fold above genomic average (**Fig. 19a**). Hence, the substantial differences in H3K9me3 that depended on the presence of TF binding sites within the interspersed repeats (Bulut-Karslioglu *et al.* 2012) was not observed here. Importantly, only 23 of the 42 described intergenic GSATs have been considered here, since the other 19 showed insufficient input coverage (i. e. <30% of the genome average input coverage) and were thus considered "unmappable". Loss of either one of the SUV39H isoforms led to a substantial decrease of H3K9me3, which was also observed for the 319 heterochromatic regions and the 179 genes that were mappable and highly methylated in wt iMEFs (**Fig. 19**). In all cases – i. e. interspersed GSATs, heterochromatic regions and genes – SUV39H2 had a greater influence on H3K9me3 compared to SUV39H1, which mirrors the effect of SUV39H2 on global and PCH H3K9me3 levels. Whereas the methylated heterochromatin regions (**Fig. 19b**) display a uniformly lower H3K9me3 level in SUV39H-deficient cells compared to wt iMEFs, the highly methylated genes showed a

bimodal distribution with a group of genes being affected by loss of SUV39H and a group of genes being SUV39H-independently trimethylated at H3K9 (**Fig. 19c**). As described elsewhere, the latter group might be regulated by G9A (Ogawa *et al.* 2002, Roopra *et al.* 2004) or SETDB1 (Ayyanathan *et al.* 2003, Schultz *et al.* 2002). The overall H3K9me3 levels of genes in iMEFs were rather low compared to the genomic average or repetitive sequences. This finding is consistent with the notion that terminal silencing in differentiated cells is governed by DNA methylation rather than histone modifications (Bird 2002, Bulut-Karslioglu *et al.* 2014).

III.3 Loss of SUV39H/H3K9me3-dependent heterochromatin is neither necessary nor sufficient for transcription activation

Despite being generally considered a silencing histone modification, the exact interplay of H3K9me3, particularly SUV39H-dependent H3K9me3, and transcription remains obscure. Here, four types of SUV39H-dependent H3K9me3 target loci were considered: PCH, interspersed GSATs, other repeat-rich interspersed heterochromatic regions (HET6387) and genes. Although H3K9me3 levels and produced RNA were mostly anticorrelated at PCH (wt, SUV39H1 sn and SUV39H2 sn cell lines, original dn cell line), PCH with high H3K9me3 in NIH3T3 cells was transcribed nonetheless. In addition, loss of PCH H3K9me3 in two SUV39H dn cell lines (E124-1 & E124-3) did not activate transcription. Similar patterns were observed at other loci. In particular, demethylated interspersed GSATs or heterochromatic regions did not get activated with the exception of two to three out of 23 GSATs (**Fig. 21a**). Interestingly, some of the regions showed basal transcription (>1 RPKM) even if highly H3K9 trimethylated in wt iMEFs (**Fig. 21a,b**) indicating that the silencing marks can be "overruled" by some activating signals. This was even more pronounced for the highly methylated protein-coding RefSeq genes, where 54 out of 119 genes (∼45%) had RNA expression values >1 RPKM. In addition, the group of genes that was demethylated in SUV39H-deficient cells (genes in the left peak in **Fig. 19c**) was not generally derepressed, but if their RNA expression changed, it mostly increased (**Fig. 21d**). Those results are consistent with the notion that SUV39H-dependent H3K9me3 stabilizes a silenced state and protects genomic loci against sporadic activation. Nevertheless, it can be "overruled" by strong and/or targeted activation (**Fig. 40**). The latter could be the case in the presence of a strong exogenous factor such as a viral activator (see below) or the activation of a signaling cascade due to extracellular signals, such as stimulation of immune response genes upon interferone signaling (Platanias 2005). One of the few genes that was strongly derepressed upon loss of one or both of the SUV39H isoforms is located within a larger (∼1 Mb) H3K9me3 domain that was SUV39H-dependent (**Fig. 22a**). Disruption of this heterochromatin domain did not significantly derepress any of the neighboring genes, suggesting again that loss of H3K9me3 was not sufficient

to activate transcription, but that an additional activating factor was required. Finally, changes in H3K9me3 levels within the gene body might not change total transcript levels. Instead, differential H3K9me3 could affect splicing as suggested elsewhere (Saint-Andre *et al.* 2011, Yearim *et al.* 2015).

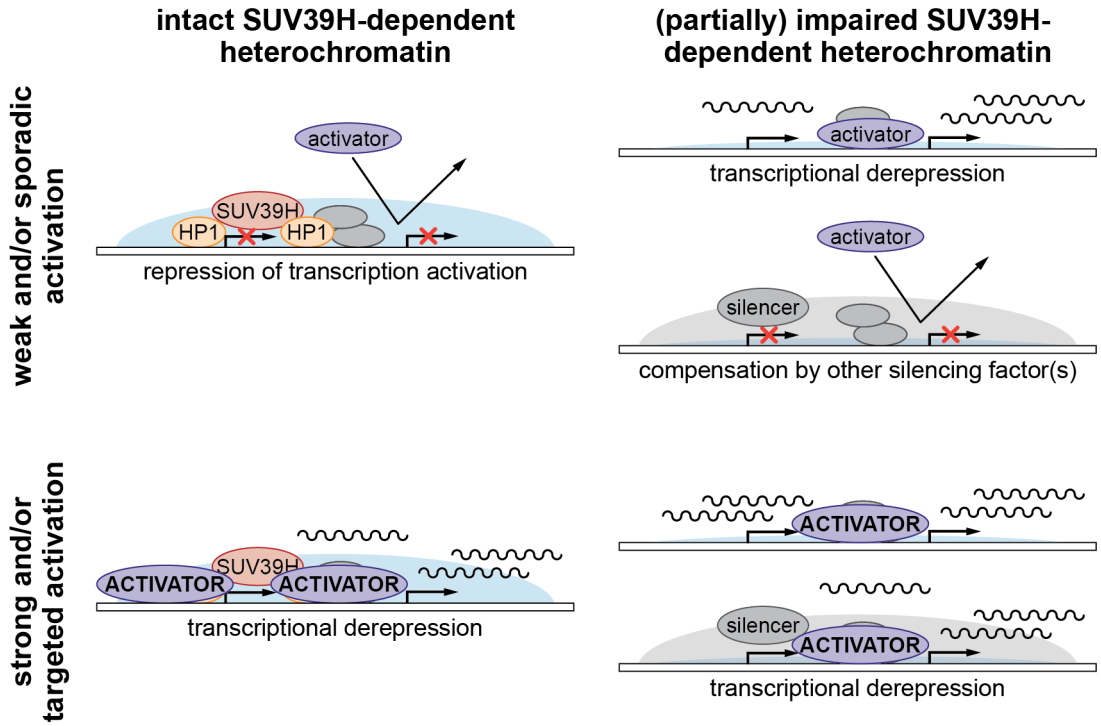

Figure 40: The role of SUV39H-dependent H3K9me3 domains in the regulation of transcription. Intact heterochromatin (left) stabilizes a transcriptionally silent state (top) by suppressing sporadic activation of genomic elements such as repeats or genes, but can be "overruled" by a strong activating signal (bottom). Loss of SUV39H2 or both SUV39H enzymes leads to partial or complete loss of SUV39H-dependent heterochromatin (right) resulting in transcriptional derepression or compensation possibly by SUV39H1 and DNA methylation or other silencing machineries (e. g. Polycomb/H3K27me3 (Saksouk *et al.* 2014)). It seems likely that impaired heterochromatin would also be readily activated in the presence of a strong or targeted activator independently of compensatory mechanisms (bottom, right). Light blue area: H3K9me3 domain; grey area: other silencing domain (e. g. H3K27me3); grey ellipses: other heterochromatin proteins (e. g. MBDs); wavy black lines: RNA transcripts.

Last, a common set of 109 genes was differentially expressed in the respective SUV39H1/2 sn and SUV39H dn cell lines as determined by DESeq2 (log2-fold change >3, $p_{adj} < 0.01$). However, only two of those genes were highly methylated at H3K9 in the first place and then demethylated upon loss of SUV39H. Thus, the observed expression changes did not directly depend on H3K9me3. Interestingly, PAX3 and PAX9 that have been implicated in targeting SUV39H-dependent heterochromatin (Bulut-Karslioglu *et al.* 2012) were among the common differentially expressed genes. Although they have been described as redundant regulators of heterochromatin, expression of PAX3 was downregulated >10-fold and

PAX9 expression was upregulated >10-fold, indicating that they are regulated differently in the context of SUV39H loss and might therefore also have unique functions.

The H3K9me3 level of the 45 commonly downregulated genes (**Fig. 22b**) tended to be higher. This is consistent with H3K9me3 being a silencing mark, but appears counterintuitive as one or two KMTs are missing in those cell lines. This could, however, be explained by overcompensation of the loss of SUV39H through other H3K9 KMTs. Contrarily, the 64 commonly upregulated genes showed no clear trend regarding their change in H3K9me3. Taken together, the differential gene expression in the different SUV39H knock-out cell lines is probably not directly related to gene body H3K9me3. A detailed analysis of H3K9me3 changes at regulatory elements – i. e. promoters and enhancers – and their relation to transcriptional activation or silencing might elucidate the functional links between loss of SUV39H and differential gene expression.

III.4 BLInCR is a rapid and reversible tool to target a variety of subnuclear structures

To test if a *bona fide* heterochromatic locus can be activated by a strong activator as suggested above (**Fig. 40**) and to analyze the kinetics of this activation, the versatile light-inducible recruitment system BLInCR was devised. Therefore, CIBN was recruited to various nuclear subcompartments, e. g. telomeres or nucleoli in the human osteosarcoma cell line U2OS 2-6-3, via fusions with compartment-specific proteins (in this case TRF1/2 or nucleolin, respectively). This led to a relocalization of PHR-tagged effectors to those CIBN-tagged compartments upon illumination with blue light (**Fig. 23**). The BLInCR system was also functional in a MEF cell line for targeting PCH foci (**Fig. 24**). So far, all proteins could be tagged with CIBN without compromising their nuclear localization, which was also reported in another study that used CIB1-TRF1 fusion constructs in HeLa cells (Choudhury *et al.* 2016). In principle, switching the optogenetic tags, i. e. tethering PHR to localizer proteins and recruiting CIBN-tagged effectors would be an intriguing modification of BLInCR, because PHR is the component that changes its confirmation upon light exposure and it would be rather immobilized. Thus, one could illuminate only part of a cell, e. g. half of the nuclear lamina, which would only target a CIBN-tagged effector to that part of the cell and create an intrinsic negative control (the unilluminated part of the cell). Unfortunately, a CIBN-tagged fluorophore could not be targeted to the nuclear lamina when cotransfecting the cells with PHR-tagged hLaminB1 fusions, irrespective whether the tag was N- or C-terminal (data not shown). This is, however, probably a lamina-specific effect, since CIB1-tagged activators have been successfully recruited to DNA-tethered CRY2 (Konermann *et al.* 2013, Pathak *et al.* 2017, Taslimi *et al.* 2016). As the lamina is a complex protein network with tight interactions between copolymerized lamin proteins (as reviewed in (Herrmann & Aebi 2004)), hLaminB1-fused

PHR might be sterically blocked which could prevent its conformational change upon illumination. In any case, the suitability of BLInCR or its modified version with switched tags would need to be assessed in the respective cell line with the respective constructs. This is facilitated by the large collection of human and mouse CIBN and PHR constructs described in this work (**Table 6**).

The recruitment kinetics of a number of effector constructs were analyzed and the characteristic time to reach half-maximal levels was in the range of 10-23 s (**Fig. 25**), which is faster than conventional inducible systems based on chemically induced dimerization. Importantly, the recruitment kinetics depend on the number of binding sites and on the concentration of the recruited molecule (Normanno *et al.* 2015). In the experiments presented here, cells expressing rather low amounts of the transfected constructs were chosen to avoid pre-recruitment of the effector in the dark. Doxycycline-induced binding of fluorescently tagged reverse TetR constructs at low concentrations has characteristic recruitment times in the order of 100 s, i. e. about an order of magnitude slower than BLInCR (Normanno *et al.* 2015). Tamoxifen- or LINuS-induced nuclear import and subsequent binding of a DBD to a reporter promoter is another order of magnitude slower with characteristic recruitment times in the order of 10-20 min (Niopek *et al.* 2014, Rafalska-Metcalf *et al.* 2010). It should be noted that LINuS and BLInCR also depend on the illumination pattern and light intensities, which were not comparable, but this is not expected to make a 100-fold difference. Recruitment of constructs containing either VP16 or hGCN5 were recruited faster than PHR-YFP control constructs. Assuming that the two globally fitted rates k_1 and k_2 correspond to the PHR-CIBN and the PHR-PHR association, VP16 and hGCN5 might limit PHR-PHR oligomerization due to their high concentration of negative charge and size, respectively. To test this hypothesis, effector concentrations could be measured to assess if the steady state amount of recruited PHR was lower for VP16- and hGCN5-containing constructs. Alternatively, an oligomerization-deficient PHR version such as CRY2 [1-507] (F507D) (Park *et al.* 2017) could be used to check if PHR-YFP control constructs would reach steady state levels faster and/or recruitment curves could be fitted with a single rate. For the purpose of transcription activation described below, the details of recruitment were rather unimportant as long as significant amounts of VP16 accumulated around the promoter within seconds, which was the case (**Fig. 25**). It should be noted that a new optogenetic system, magnets, has been shown to achieve accumulation with characteristic recruitment times as low as 1.5 s (Kawano *et al.* 2015). Last, the system was reversible on the time scale of minutes ($t_{1/2}$ =∼5 min) with no effector detectable after 15-20 min (**Fig. 26**). This is an intermediate half-life time compared to other optogenetic systems such as magnets ($t_{1/2}$ =∼7 s, (Kawano *et al.* 2015)) or PHYB that remains in its active confirmation for hours after induction if not reversed by illumination with a far-red light pulse (Ni *et al.* 1999). For analyzing the activation kinetics of a reporter locus as conducted here, the time scale of reversibility in

the order of minutes was suitable as constant presence of significant amounts of activator at the array could be achieved with imaging every 2 to 5 minutes, which was acceptable with respect to laser-induced phototoxicity and photobleaching.

The main limitations of BLInCR are (i) the use of fusion constructs that might impair localizer or effector function, (ii) the level of transfected PHR-effector constructs, which can cause leakiness in the dark if too high and, (iii) the transfected effector constructs that are present (and possibly active) before illumination. The first is an issue of fusion constructs in general and thus has to be tested for each BLInCR design. The second is not problematic for single live cell experiments, because cells with appropriate transfection levels can be selected. For bulk experiments, the generation of stable cell lines with low expression values should be considered. The third point strongly depends on the effector: VP16 has been described to only activate genes to whose promoter it is specifically targeted to (Polstein *et al.* 2015). In other cases such as targeting the FokI DNA endonuclease to telomeres (as in (Tang *et al.* 2013) via fusion to TRF1), this might be more problematic since freely available PHR-FokI would cut in the proximity of any GGATG sequence in the genome (Sugisaki & Kanazawa 1981). This limitation might be overcome by using an additional blue light-dependent optogenetic approach that inactivates the effector, e. g. by caging a functional peptide using LOV domain-dependent TULIPs (Strickland *et al.* 2012). In summary, BLInCR is a versatile, rapid and reversible tool to increase local effector concentrations that has some inherent limitations due to the PHR/CIBN system used. However, those limitations can be measured and controlled for and possibly also improved by implementing recently published modifications of the optogenetic components (Taslimi *et al.* 2016).

III.5 BLInCR is a suitable tool to monitor transcription in living cells

BLInCR of the viral transcription activator VP16 to the promoter-proximal *tet*O array in the U2OS 2-6-3 reporter cell line (**Fig. 27a**) efficiently activated transcription (**Fig. 27d**). Light-dependent nuclear clearance of optogenetic constructs as reported for PHR-DBD constructs in another reporter system (Pathak *et al.* 2017) was not observed here. Activation by VP16 BLInCR worked robustly, but resulted in lower accumulated RNA levels (**Fig. 28d**) and in less pronounced decondensation of the reporter array after constitutive activation (**Fig. 32**) compared to VP16 recruitment via TetR fusions. Since light-dependent activation has been shown to depend on the integrated blue light dose (Kennedy *et al.* 2010), BLInCR-dependent transcription activation as well as the associated chromatin structure changes could probably be tuned to mimic endogenous decondensation (Boettiger *et al.* 2016). One could speculate that the observed difference between activation with BLInCR or fusion constructs might be related to the VP16-DNA

distance. This distance is probably smaller in the case of fusion constructs. In addition, it would be interesting to see if the transcriptional output can be modulated by the light intensity as described for the number of light pulses (Kennedy *et al.* 2010) or for two different light intensities for yeast (Pathak *et al.* 2014). Apart from the difference in transcriptional output that was due to the recruitment method (BLInCR vs. fusion), less MS2 RNA was detected when targeting VP16 to *lac*O compared to *tet*O. This was probably due to the distance of *lac*O to the minimal promoter (~4,000-13,000 bp), since all effects of VP16 – i. e. recruiting activating factors (Hirai *et al.* 2010) or perturbing local nucleosome structure (Erkina & Erkine 2016) – are rather short-distance effects.

Upon recruitment of VP16, GFP-tagged labeled pol II readily accumulated at the reporter array (**Fig. 27c**) and IF staining for active pol II marks, i. e. phosphorylation of serines 2 & 5, was also observed albeit at different levels. The initiation-associated pSer5 mark was heavily enriched in all activated cells. In contrast, the elongation-associated pSer2 mark was far less enriched and not detectable in most cells (**Fig. 30**). Hence, the pSer2 mark was not necessary for the release of paused pol II or the pSer2 levels were below the detection limit. The latter scenario also seems plausible since pSer2 is generally much more enriched towards the 3' end of transcribed genes (Heidemann *et al.* 2013). A single reporter gene might be too short (~3,000 bp) compared to most human genes with an average size >10,000 bp (Grishkevich & Yanai 2014) for detectable pSer2 enrichment. To address this question in more detail, the specific CTD kinases including CDK7 (pSer5) as well as CDK9 and CDK12 (pSer2) could in principle also be detected in cells activated by VP16 BLInCR.

Detection of MS2 RNA via fluorescently labeled MCP is a widely used tool to visualize on-going transcription in living cells (Darzacq *et al.* 2009, Larson *et al.* 2013) and was also compatible with BLInCR (**Fig. 27b**). Even though the enrichment of MCP at the array is not suitable to detect absolute RNA levels, it does represent the cell-to-cell heterogeneity of RNA accumulation measured by RNA FISH (**Fig. 29a,b**). For more comparability between cells, a stable cell line with more homogeneous MCP expression levels would be helpful. In addition, one could calibrate the measured intensity to absolute reader particle numbers using fluorescence correlation spectroscopy (FCS) similar to previous studies characterizing MCP binding (Wu *et al.* 2012). However, to analyze the kinetics of transcription activation in single cells, the relative enrichment of MCP over time is also an informative measure. To read out endogenous RNA species such as GSATs (**Fig. 27d,e**), more elaborate detection methods such as molecular beacons (Tyagi & Kramer 1996) or sticky-flares (Briley *et al.* 2015) are required. For proof-of-concept experiments, RNA quantification by qRT-PCR is, however, also suitable when high transfection efficiencies are reached. In summary, recruiting a strong viral transcription activator such as VP16 to the proximity of a target promoter via BLInCR allows robust induction of pol II-dependent transcription. RNA detection via MCP is also suitable for MS2 loop-containing

transcripts, but can also be realized by PCR-based methods (**Fig. 27e,28a**). In addition, coinciding recruitment of other fluorescently tagged effectors such as pol II and RNA accumulation can also be monitored in living cells.

III.5.1 Transcription-associated changes of the local chromatin environment can be traced by fluorescently tagged reader domains

After constitutive activation of transcription, changes in the modification state of H3K9 were minor, with H3K9ac levels increasing slightly and H3K9me3 levels remaining virtually unchanged (**Fig. 31**). Particularly the subtle changes in H3K9ac are in contrast to the strong hyperacetylation at many histone tail residues observed by others (Rafalska-Metcalf *et al.* 2010, Tumbar *et al.* 1999). Interestingly, Tumbar and colleagues did not see any hyperacetylation of H3K9 in ∼30% of the cells and only a weak effect in another ∼35%, indicating that strong hyperacetylation of H3K9 is not a universal feature. Since BLInCR is particularly useful to monitor the consequences of targeting an effector to a specific locus in living cells over time, fluorescently tagged reader proteins or protein domains are of particular interest as they can also be used in live cell experiments, e. g. in addition to the RNA read out. The bromodomain-containing proteins 2 and 4 (BRD2/4) are particularly interesting candidates as their bromodomains can bind acetylated histones (LeRoy *et al.* 2008, Vollmuth & Geyer 2010) and BRD4 is also involved in facilitating the transition of paused pol II to productive elongation (Jonkers & Lis 2015). They could therefore serve as live cell readers for acetylated histones and possibly as markers for ongoing transcription. However, both were not recruited to the reporter array activated by VP16 BLInCR in U2OS 2-6-3 cells (**Fig. 33**). BRD4 has been shown to be dispensable in some cases of transcription initiation via a viral activator (Yang *et al.* 2005), although it is essential for the activation of stimulus-dependent endogenous genes (Patel *et al.* 2013). If the observed absence of BRDs from activated arrays is related to the viral nature of activation used in this system, it remains confusing why the BRDs are recruited when activating transcription using TetR/LacI-VP16 fusion constructs (**Fig. 33**), which has also been previously described elsewhere (Rafalska-Metcalf *et al.* 2010). How independent acetylation binding and necessary coactivation functions of BRDs at the array are interrelated is currently an open question. It would be interesting to see if arrays activated by fusion constructs are enriched for histone acetylation and if they could also be activated in the absence of BRDs. However, for the advancement of the BLInCR methodology for live cell monitoring of processes associated with transcription activation, it might be more feasible to consider fluorescently tagged CDK7 and/or CDK9. The latter factors set the possibly indispensable phosphorylation marks on the pol II CTD, and could be used as tracers of actively ongoing transcription as used for live cell imaging of transcription factories (Ghamari *et al.* 2013).

H3K9me3 levels at the reporter array were not affected by transcription activation

as determined from IF staining (**Fig. 31**). Nevertheless, live cell reader domains for H3K9me3, namely the mouse chromodomain of HP1β (mCD) and the human ADD domain (hADD), were tested on inactive arrays (**Fig. 34a**). Accumulation of mCD was more pronounced than hADD accumulation, indicating that the mCD might be a more suitable H3K9me3 reader domain. Notably, both mCD and hADD showed less accumulation at the array when an array marker (i. e. fluorescently tagged TetR or LacI) was tethered to the *tet*O repeats compared to the *lac*O repeats. DNA binding of TetR and/or LacI could interfere with H3K9me3 recognition by reader domains. In that case, the observation would indicate that histones methylated at H3K9me3 are predominantly found at the *tet*O repeats close to the promoter. ChIP of H3K9me3 and subsequent qPCR or DNA sequencing could shed light on the distribution of H3K9me3 within this silenced locus and might explain the interdependencies of array labeling and histone modification read out. Importantly, some cells with activated arrays as judged by CFP-SKL expression and/or MCP accumulation showed enrichment of the mCD, although the average H3K9me3 level was lower in activated cells compared to inactivated ones (**Fig. 34**). This is consistent with the notion that H3K9me3 is generally a silencing mark, but cannot prevent transcription in the presence of a strong, targeted activator (as in **Fig. 40c**, left). The H3K9 modifications monitored here displayed only subtle changes after 24 h of activation. Hence, tracing those changes at the array in real-time experiments during activation would not be very insightful here. However, it is in principle possible and recruitment of different (non-viral) activators (e. g. STAT transcription factors) or use of other reader domains might provide useful to dissect the timing of histone modification changes in the context of transcription activation.

III.5.2 BLInCR adds fast reversibility to the toolbox of live cell transcription activation methods

Targeting VP16 to tetO in U2OS 2-6-3 cells via BLInCR resulted in RNA accumulation within minutes after first exposure to blue light and dispersion of the signal after dissociation of the effector in the dark (**Fig. 37b,c**). A second VP16 recruitment by blue light resulted in another increase of RNA accumulation, showing that BLInCR can be used to reversibly trigger effector relocalization within minutes. Light-dependent reversibility is easier to use than wash-out of e. g. doxycycline and can be controlled more precisely. The use of far-red fluorescent proteins such as iRFP713 (Filonov *et al.* 2011) in combination with established red fluorophores such as tagRFP allowed to quantify both effector and reader accumulation while avoiding renewed recruitment. Thus, BLInCR can be used to apply tailored illumination patterns resulting in well-controlled effector accumulation and dissociation. Combining this methodology with live cell readers for RNA and histone modifications, BLInCR is a valuable tool to study epigenetic memory of transcription states.

III.6 The time course of transcription activation resolved by BLInCR reveals a biphasic activation kinetics

First transcripts accumulated rapidly after triggering VP16 accumulation at the array by BLInCR and reached detectable amounts in less than two minutes in some cells (**Fig. 36a,b**). The subsequent lag phase that displayed rather unchanged amounts of accumulated RNA at the array was followed by another rapid RNA accumulation phase. The rapid response to the activator of at least some cells in the untreated case and all cells in SAHA-treated cells is consistent with the idea of a fraction of poised reporter genes that can be rapidly activated. Histone hyperacetylation after SAHA treatment could therefore increase this fraction of reporter copies that is poised or primed for activation, which seems plausible as H4ac is increased upon SAHA treatment (**Fig. 35a**) and might recruit some of the PIC components such as TFIID that has been shown to interact with acetylated lysine residues on the H4 tail (Jacobson *et al.* 2000). During the lag phase, another – inactive or weakly active – fraction of reporters might undergo a maturation process allowing its delayed activation and consequently increasing levels of RNA at the array after twenty to forty minutes (**Fig. 36a,b**). Processes occurring during this maturation can be diverse. They range from delayed recruitment of coactivators or nucleosome remodeling (Nocetti & Whitehouse 2016, Venkatesh & Workman 2015) to interaction with neighboring actively transcribing reporter genes (Li *et al.* 2012, Papantonis & Cook 2013) or relocalization (Therizols *et al.* 2014) within the array. The array size did not change during this initial activation phase (**Fig. 36b**). Thus, massive decondensation on the scale observable by light microscopy plays no role during this initial activation phase, but might rather be a downstream effect of transcription. Similarly, extensive changes in histone modifications would also not be expected, since they were not observed after constitutive activation over night (**Fig. 31, 34b**). Transcription was not initiated homogeneously across the reporter array (**Fig. 36a**). Accordingly, the notion that spatially distinct compartments within the array exist and need to contact or spread to other loci also seems reasonable. Promoters functioning as specific enhancers for other genes (Dao *et al.* 2017, Diao *et al.* 2017) as well as promoter-promoter interactions within a gene cluster have been described recently (Corrales *et al.* 2017). These could contribute to enhanced activation of repressed genes. As for the corecruitment of effectors, BLInCR is well-suited tool to monitor the recruitment kinetics of components that are not directly recruited by BLInCR (e. g. GFP-RPB1 as in **Fig. 36c**). Notably, simulations revealed that rapid recruitment as achieved by BLInCR and time-resolved detection of RNA as realized by RNA detection in living cells is required to resolve the biphasic kinetics of transcription activation (Rademacher *et al.* 2017).

Assessing the persistence of the actively transcribing state proved difficult in live cells due to extensive bleaching of iRFP713 and heterogeneous responses across single cells

(**Fig. 37b,c**). However, an initial experiment using qRT-PCR as a read out showed that RNA levels decayed after removing the light trigger (**Fig. 37d**). This finding indicates that the actively transcribing state cannot be maintained in the absence of the activating effector. Other studies that have probed the persistence of the transcriptionally silent state reported varying results. In some cases, silencing modifications depended directly on the presence of an effector (Kungulovski *et al.* 2015) or no memory was observed after mitotic division (Brueckner *et al.* 2016). This is in contrast to induced silent states that can display prolonged persistence for hours or days without the initial trigger (Bintu *et al.* 2016, Hathaway *et al.* 2012).

Taken together, even a strong activator such as VP16 cannot activate a gene array in a single step, but requires some maturation or rearrangement to result in enhanced activation (**Fig. 41**). In addition, rapid initial activation was modulated by hyperacetylation of histones after treatment with the histone deacetylase inhibitor SAHA. Finally, transcription activity was decreased or shut down when the effector was no longer recruited to the reporter.

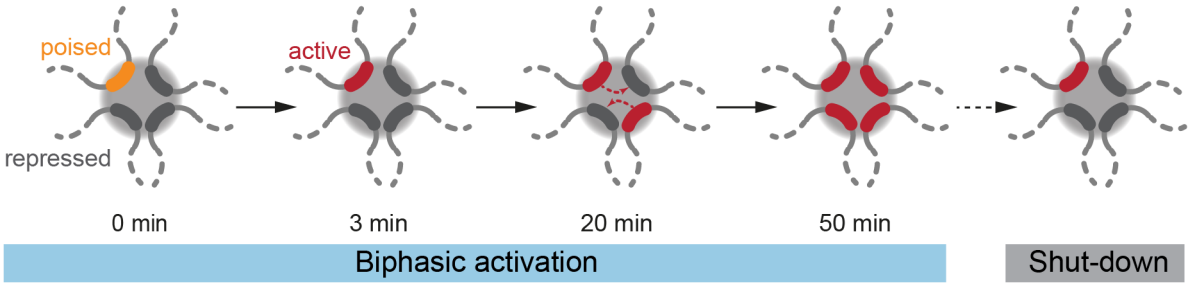

Figure 41: Biphasic activation model of a gene array. In the initial rapid activation phase, a fraction of poised or primed reporters gets activated producing detectable RNA amounts within ~2 min. During the maturation phase, repressed promoters get activated leading to enhanced transcription after 20-40 min. Upon loss of the activating effector in the dark, transcription activity ceases leading to a gradual decrease of total RNA. Figure adapted from (Rademacher *et al.* 2017).

III.7 Conclusion and perspectives

The mechanisms by which the genome compartmentalizes into distinct functional units within the nucleus after cell division are difficult to dissect experimentally and might vary for different compartments. In this study, the specificity and memory of the SUV39H/H3K9me3-dependent heterochromatic state was analyzed. The results obtained here showed: i) that the H3K9me3 'writer' enzyme SUV39H1 remains chromatin-bound throughout the cell cycle and ii) that the mobility of the methyl-DNA-binding protein MECP2 was independent of SUV39H. Therefore, SUV39H1 might 'bookmark' heterochromatic sites throughout cell division and contribute to targeting other heterochromatin factors to those sites after mi-

tosis. In addition, MECP2 could also be a factor that actively guides SUV39H enzymes to their sites of action. In this case, the pressing questions are, how MECP2 specificity can be explained and if MECP2 is another bookmarking factor that might also contribute to SUV39H binding to chromatin during mitosis. In established SUV39H-centered domains such as PCH, both isoforms of the histone methyltransferase are enriched, which might confer stability to the heterochromatic state of the domain. Analysis of SUV39H1/2 single null iMEFs revealed that the function of both isoforms is not completely redundant and that SUV39H2 contributes more to H3K9me3 at PCH. The same trend was also observed in a genome-wide analysis of interspersed heterochromatic repeats and genes. However, SUV39H-dependent H3K9me3 did not always repress transcription and its loss only sporadically led to derepression of the respective regions. This raised the question how the activation of a heterochromatic locus proceeds. Therefore, the activation of a bona fide heterochromatic reporter locus in the U2OS 2-6-3 cell line was dissected in single living cells. To be able to precisely trigger activation, the Blue Light-Induced Chromatin Recruitment (BLInCR) tool was devised. It allowed to trigger the relocalization of the strong viral activator VP16 to the reporter locus by light. Using this technique, transcription was rapidly activated, showing that the heterochromatic state can be overwritten by a strong activator and that activation was upstream of histone modifications or large-scale chromatin rearrangements. VP16 targeting by BLInCR also revealed that activation of a heterochromatic gene cluster was biphasic and that histone acetylation influenced the shape of the activation curve. BLInCR can also be used to disturb the epigenetic landscape at genomic loci by targeting epigenetic modifiers such as demethylases or acetyltransferases to further elucidate the role of histone modifications in silencing and activation.

In summary, mechanistic details on specificity and memory of endogenous SUV39H-dependent silencing and the contribution of the different SUV39H isoforms were dissected here. In combination with the analysis of the light-induced activation kinetics of a heterochromatic locus, those results provide insight into the establishment and maintenance as well as on the stability of this chromatin state. I envision that understanding the underlying principles of silencing and activating heterochromatin will be instrumental in the identification of drug targets and the development of epigenetic drugs, because epigenetic deregulation and aberrant transcription of repetitive sequences are frequently observed in cancer and other diseases.

IV Materials & Methods

IV.1 Materials

Table 6: Plasmids constructed within my thesis work.

construct name	CloneX ID	construct name	CloneX ID
CIBN constructs		**PHR constructs**	
CIBN-tagBFP-hLaminB1	776	PHR-EYFP	835
CIBN-EYFP-hLaminB1	816	PHR-EYFP-NLS	986
CIBN-tagBFP-hLaminB1-IRES2	821	PHR-mCherry-NLS	1006
CIBN-tagBFP-hLaminB1-IRES2-mCherry	825	PHR-EYFP-VP16	836
CIBN-tagBFP-hLaminB1-IRES2-GBP-PHR	831	PHR-mCherry-VP16	948
CIBN-tagBFP-TetR	929	PHR-EYFP-hPMLIII	937
CIBN-EYFP-TetR	930	PHR-mCherry-hPMLIII	942
CIBN-TetR	865	PHR-EYFP-CDK1AF	962
CIBN-TetR-tagRFP-T	875	PHR-mCherry-CDK1AF	963
CIBN-tagBFP-hTRF1	943	PHR-EYFP-hGCN5	1003
CIBN-EYFP-hTRF1	944	PHR-EYFP-hGCN5mut	1004
CIBN-hTRF1-tagRFP-T	945	PHR-EYFP-hGCN5-NLS	1005
CIBN-tagBFP-hTRF2	931	PHR-EYFP-NBS1	1007
CIBN-EYFP-hTRF2	932	PHR-mCherry-NBS1	1008
CIBN-hTRF2-tagRFP-T	946	PHR-EYFP-hHP1β	1011
CIBN-tagBFP-hNCL	933	PHR-mCherry-hHP1β	1010
CIBN-EYFP-hNCL	934	hLaminB1-PHR[a)]	797
CIBN-tagBFP-hPMLIII	935	PHR-tagBFP-hLaminB1[a)]	817
CIBN-EYFP-hPMLIII	936		
CIBN-EYFP-mHP1β	1013	**reader domains**	
CIBN-mSUV39H1	1089	mCherry-mBrd2	964
CIBN-EYFP-mSUV39H1	1086	mKate2-mBrd2	869
CIBN-EYFP-mSUV4-20H2	1087	mCherry-MS2	876
CIBN-mCherry	815	CFP-MS2	947
		tagBFP-MS2	941
fluorescently tagged effectors		hADD-mKate2	926
EGFP-mSUV39H1-ΔN89	698	hADD-CFP	927
mSUV39H1-ΔN89-EGFP	699	hADD-tagBFP	928
mSUV39H1-ΔN89-tagRFP	700	mCD-CFP	883
Flag-mSUV39H1-ΔN89	701	mCD-tagBFP	882
EGFP-mSUV39H2	789	mCherry-mB2BD[b)]	866
EGFP-CDK1AF	961	mKate2-mB2BD[b)]	870
EGFP-CyclinBΔaa1-166	1009	mCherry-mBRD4[b)]	867
EYFP-hLaminB1	790	mKate2-mBRD4[b)]	871
		mCherry-mB4BD[b)]	868
other LacI/TetR fusion proteins			
hNCL-EGFP-LacI	949		
hNCL-CFP-LacI	950		
tagBFP-LacI-VP16	775		
tagBFP-TetR-VP16	873		
GFP-TetR-VP16	874		

[a)] These constructs did not work as they did not lead to CIBN-mKate2-NLS accumulation upon illumination with blue light. [b)] These constructs did not work in the sense that they did not accumulate when co-transfected with any VP16 construct.

Table 7: Other plasmids used.

construct name	CloneX ID	source/comment
PAX3-GFP	645	(Bulut-Karslioglu *et al.* 2012)
PAX5-GFP	646	(Bulut-Karslioglu *et al.* 2012)
PAX7-GFP	647	(Bulut-Karslioglu *et al.* 2012)
PAX9-GFP	648	(Bulut-Karslioglu *et al.* 2012)
pCIBN(deltaNLS)-pmGFP	795	Addgene #26867 (Kennedy *et al.* 2010), cloning of CIBN
pCRY2PHR-mCherry-N1	796	Addgene #26866 (Kennedy *et al.* 2010), cloning of PHR
pGEX-GP-2-ATRX ADD	851	Addgene #59698/A. Jeltsch lab, cloning of hADD
pAdEasy Flag GCN5	854	Addgene #14106 (Lerin *et al.* 2006), cloning of human GCN5 constructs
pAd-Track Flag GCN5 Y621A/P622/A	855	Addgene #14425 (Lerin *et al.* 2006), cloning of human GCN5mut
GFP-RPB1	1242	largest pol II subunit for live cell imaging (Sugaya *et al.* 2000), from Jan Ellenberg (EMBL)
YFP-mBRD2	862	as a reader and for cloning of other Brd2 constructs (Rafalska-Metcalf *et al.* 2010)
YFP-mBRD4	863	as a reader and for cloning of other Brd4 constructs (Rafalska-Metcalf *et al.* 2010)
GFP-MECP2	597	(Schmiedeberg *et al.* 2009)
pEGFP-MS2_coat	547	J.-Philipp Mallm; cloning of MS2
GBP-TetR	809	F. Erdel; cloning of TetR
GBP-PHR	830	F. Erdel; cloning of GBP-PHR IRES2 constructs
PHR-hLaminB1	777	F. Erdel; cloning of PHR-hLaminB1 constructs
CIBN-EYFP-hLaminB1	816	F. Erdel; cloning of hLaminB1
CIBN-EYFP-hSUV39H1	1085	F. Erdel
CIBN-LacI	828	F. Erdel
CIBN-mKate2-NLS	791	F. Erdel
mCD-mKate2	628	F. Erdel; chromodomain of mouse HP1β, H3K9me3 reader (Muller-Ott *et al.* 2014) and cloning of other mCD constructs
CIBN-LacI-tagRFP-T	853	K. Mueller-Ott/Ronja Rappold; cloning of CIBN-. . .-tagRFP-T
PHR-iRFP713	991	J. Trojanowski
PHR-iRFP713-VP16	994	J. Trojanowski
GFP-LacI-VP16	264	I. Chung; cloning of VP16 and LacI
EGFP-N1 mSUV39H1	527	K. Mueller-Ott/C. Bauer; cloning of mSuv39h1-ΔN89
SUV39H2-EGFP	440	Schotta lab/ K. Mueller-Ott; cloning of SUV39H2
CMeGFP-mSUV39H1	432	C. Bauer, cloning of mSUV39H1
CMeGFP-mouseHP1β	420	C. Bauer, cloning of mHP1β
pTRE3G-ZsGreen	718	K. Deeg; cloning of IRES2
EGFP-C1-TRF2	231	T. Jegou (Jegou *et al.* 2009); cloning of hTRF2
EGFP-C1-PMLIII	273	T. Jegou (Jegou *et al.* 2009); cloning of hPMIII
CFP-LacI	239	T. Jegou
RFP-NCL	897	M. Caudron-Herger (Caudron-Herger *et al.* 2015); cloning of hNCL
pEGFP-N1_TRF1	837	D. Braun/I. Chung (Chung *et al.* 2011); cloning of hTRF1
p2GFP-NBS1	576	I. Chung/(Lukas *et al.* 2003), cloning of NBS1
ECFP-HP1β	558	I. Chung, cloning of hHP1β
CDK1AF-containing plasmid	-	from Pierre Wehler (Barbara Diventura group)
tagBFP-containing plasmid	-	from Ashish Goyal (Sven Diederichs devision)
ENTR221-CyclinB	1012	DKFZ core facility (clone 14754003)
SUV4-20H2-GFP	505	

Table 8: Cell lines used within this thesis.

cell line	CloneX ID	source/comment
NIH3T3	370	commercial (DSMZ)
NIH3T3	675	(Muller *et al.* 2009)
iMEF wild type	669	(Peters *et al.* 2001)
iMEF SUV39H double null	668	(Peters *et al.* 2001)
NIH3T3 SUV39H1-GFP	369	Thomas Jenuwein (MPI Freiburg), doxycyline-inducible SUV39H1-GFP expression
iMEF E469-2	706	Gunnar Schotta (LMU Munich), SUV39H1 single null
iMEF E469-3	707	Gunnar Schotta (LMU Munich), SUV39H1 single null
iMEF E471-2	708	Gunnar Schotta (LMU Munich), SUV39H2 single null
iMEF E480-2	710	Gunnar Schotta (LMU Munich), SUV39H2 single null
iMEF E480-6	709	Gunnar Schotta (LMU Munich), SUV39H2 single null
iMEF E124-1	807	Gunnar Schotta (LMU Munich), SUV39H double null
iMEF E124-3	808	Gunnar Schotta (LMU Munich), SUV39H double null
mESC wt26	804	Gunnar Schotta (LMU Munich), ESC wild type control (original, with FBS)
U2OS 2-6-3	809	David Spector & Susan Janicki (Janicki *et al.* 2004)

Table 9: Antibodies used within this thesis.

antigen/specification	host	dilution/ amount	source/comment
H3K9me3	rabbit	1:500 (IF) 1:1000 (WB) 4 µg (ChIP)	Abcam #8898
H3	rabbit	4 µg (ChIP)	Abcam #1791
normal IgG	rabbit	4 µg (ChIP)	R&D Systems #AB-105-C
H3K9me3	rabbit	1:1000 (WB)	ActiveMotif #39161
H3K9ac	rabbit	1:1000 (WB)	ActiveMotif #39137
H3K27ac	rabbit	1:1000 (WB)	Abcam #4729
H4ac	rabbit	1:2000 (WB)	Millipore #06-866
H3	mouse	1:1000 (WB)	Cell Signaling Technology #14269
rabbit IgG, HRP-linked	goat	1:2000 (WB)	Cell Signaling Technology #7074
mouse IgG, HRP-linked	horse	1:2000 (WB)	Cell Signaling Technology #7076
HP1α	mouse	1:300 (IF)	Euromedex (2HP-1H5-AS)
H3K9ac	rabbit	1:500 (IF)	ActiveMotif #39917
GFP	rabbit	1:500 (IF)	Abcam #290
RNA pol II CTD phospho Ser2	rat	1:100 (IF)	clone 3E10, ActiveMotif #61083
RNA pol II CTD phospho Ser5	rat	1:100 (IF)	clone 3E8, ActiveMotif #61085
mouse IgG, Alexa Fluor 488-linked	goat	1:500 (IF)	Life Technologies #A11029
rabbit IgG, Alexa Fluor 568-linked	goat	1:500 (IF)	Life Technologies #A11036
rat IgG, Alexa Fluor 647-linked	goat	1:500 (IF)	Life Technologies #A21247

IF: immunofluorescence, WB: western blot, ChIP: chromatin immunoprecipitation.

Table 10: Buffer composition.

buffer	application	composition	pH
phosphate buffered saline (PBS)	universal	1.7 mM KH_2PO_4, 137 mM NaCl, 10 mM Na_2HPO_4, 2.7 mM KCl	7.2
Tris EDTA (TE)	universal	10 mM Tris, 1 mM EDTA	8.0
Tris base, acetic acid, EDTA (TAE)	cloning	40 mM Tris, 13.8 mM $C_2H_4O_2$, 1 mM EDTA, pH 8.0	
Tris buffered saline (TBS)	WB	50 mM Tris-HCl, 150 mM NaCl	7.5
SDS running buffer	WB	25 mM Tris, 192 mM glycine, 0.1% w/v sodium dodecyl sulfate (SDS)	
transfer buffer	WB	25 mM Tris, 192 mM glycine, 0.5% w/v SDS, 20% v/v methanol (added before use)	
CSK buffer	RNA-FISH	100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 10 mM piperazine-N,N'-bis(2-ethanesulfonic acid) (PIPES), 0.5% Triton X-100	
saline-sodium citrate (SSC)	RNA-FISH	150 mM NaCl, 15 mM $Na_3C_6H_5O_7$ (trisodium citrate)	7.0
hybridization buffer (2×)	RNA-FISH	4× SSC, 20% dextrane sulfate, 2 mg/ml bovine serum albumin (BSA)	
swelling buffer	ChIP-seq	25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Hepes) pH7.8, 1 mM $MgCl_2$, 10 mM KCl, 0.1% NP-40, 1 mM Dithiothreitol (DTT), COMPLETE protease inhibitor (Roche, Germany), 0.5 mM phenylmethylsulfonyl fluoride (PMSF)	7.8
MNase buffer	ChIP-seq	25 mM KCl, 4 mM $MgCl_2$, 1 mM $CaCl_2$, 1× protease inhibitor (Cell Signalling, UK), 50 mM Tris/HCl pH7.4	7.4
sonication buffer (1×)	ChIP-seq	10 mM Tris pH8.0, 200 mM NaCl, 1 mM EDTA, 0.5% N-lauroylsarcosine, 0.1% Na-deoxycholate	8.0
elution buffer	ChIP-seq	50 mM Tris pH8.0, 1 mM EDTA, 1% SDS, 50 mM $NaHCO_3$	8.0
high-salt buffer	ChIP-seq	50 mM Hepes pH7.9, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS	7.9
Li buffer	ChIP-seq	20 mM Tris pH8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate	8.0
iT buffer	RNA-seq	25 mM Tris-HCl, 137 mM NaCl, 5 mM KCl, 0.3 mM Na_2HPO_4, 0.5 mM $MgCl_2$, 0.7 mM $CaCl_2$	7.5

Table 11: Primers and probes used in this thesis.

probe/primer	application	remarks	sequence
Atto565-labeled MS2 probe[a)]	RNA-FISH	5'-label	5'-GTC GAC CTG CAG ACA TGG GTG ATC CTC ATG TTT TCT AGG CAA TTA-3' (Goodier *et al.* 2010)
human actin primer	qRT-PCR	length of main fragment: 194 bp	fw: 5'-TCC CTG GAG AAG AGC TAC GA-3'
			rv: 5'-AGC ACT GTG TTG GCG TAC AG-3'
CFP/SKL primer[b)]	qRT-PCR	length of main fragment: 104 bp	fw: 5'-GTC CGG ACT CAG ATC TCG A-3'
			rv: 5'-TTC AAA GCT TGG ACT GCA GG-3'
mouse GSAT primer	qRT-PCR	length of main fragment: 126 bp	fw: 5'-TGG CGA GAA AAC TGA AAA TCA CG-3'
			rv: 5'-TCT TGC CAT ATT CCA CGT CCT AC-3'
mouse actin primer	qRT-PCR	length of main fragment: 133 bp	fw: 5'-TAT CCT GAC CCT GAA GTA CC-3'
			rv: 5'-CTC GGT GAG CAG CAC AGG G-3'
mouse GAPDH primer	qRT-PCR	length of main fragment: 125 bp	fw: 5'-TAT GTC GTG GAG TCT ACT GG-3'
			rv: 5'-ACA CCC ATC ACA AAC ATG GG-3'

[a)] from T. Pankert; [b)] designed by J. Trojanowski.

IV.2 Methods

IV.2.1 Molecular cloning

Within this thesis, I cloned 42 optogenetic constructs, 15 fluorescently tagged reader domains and 13 other fluorescently tagged effector constructs. All of those plasmids are listed in **Table 6**. Details on the cloning strategy for each construct as well as plasmid maps can be found in the AG Rippe CloneX database with the associated reference number provided in **Table 6**. In principle, appropriate inserts were amplified from the plasmids listed in **Table 7** by gradient polymerase chain reaction (PCR) with the following general program: 3 min at 98°C (step 1, initialization), 10 s at 98°C (step 2, denaturation), 30 s at 50-65°C (step 3, annealing, temperature gradient was generally used), 30-150 s at 72°C (step 4, elongation, time depended on the insert size), 2 min at 72°C (step 5, final elongation) and going from step 4 to step 2 34× for a total of 35 cycles. For one reaction, ∼40-200 ng of template were mixed with 0.5 µM forward primer, 0.5 µM reverse primer, 25 mM deoxynucleoside triphosphate (dNTPs) mixture (ThermoFisher Scientific, USA), Q5 polymerase reaction buffer (5x) and 2 U Q5 polymerase (New England Biolabs, USA). PCR products were purified using the Wizard PCR cleanup system (Promega, USA) and their sizes were verified by electrophoresis using 1% agarose (LE agarose, Biozym, Germany)/TAE gels. Inserts and backbones were cut with the appropriate restriction enzymes (see database entries for details) according to the manufacturer's protocols (ThermoFisher Scientific or New England Biolabs, USA). Fragments were separated on 1% cloning agarose (LE genetic pure agarose, Biozym, Germany)/TAE gels, bands were cut out and DNA fragments were isolated using the Nucleospin gel cleanup system (Macherey Nagel, Germany) or DNA gel extraction spin columns (Freeze 'n squeeze, Bio-Rad Laboratories, USA). Fragments were ligated using T4 ligase (ThermoFisher Scientific, USA) or Quick ligase (New England Biolabs, USA) and competent bacteria (XL10 or $dam^{-/-}/dcm^{-/-}$ JM110, Agilent Technologies, USA) were transformed with the plasmids, cultured in LB medium (1% w/v tryptone, 0.5% w/v yeast extract, 86 mM NaCl, pH7.0) at 37°C on a shaker over night and plasmid DNA was isolated using the Nucleospin plasmid purification system (Macherey Nagel, Germany). Isolated DNA was test digested with appropriate restriction enzymes and the inserts of promising clones were sequenced.

IV.2.2 Cell culture and transfection

Cell lines used within this thesis are listed in **Table 8**. NIH3T3 and iMEF cells were cultured in high glucose (4.5 g/l) Dulbecco's Modified Eagle Medium (DMEM, #31053, ThermoFisher Scientific Inc., USA) supplemented with 10% fetal bovine serum (FBS, PAN Biotech GmbH, Germany), 2 mM stable glutamine (PAN Biotech GmbH, Germany), 1x Penicillin/Streptomycin (PAN Biotech GmbH, Germany) at 37°C and 5% CO_2. To induce stable expression of SUV39H1-GFP in the respective cell line (**Table 8**), 3 µg/ml doxycy-

cline (Sigma-Aldrich, USA) were added to the culture medium. U2OS 2-6-3 were cultured in low glucose (1.0 g/l) DMEM (#11880, ThermoFisher Scientific Inc., USA) with the same supplements as above except for the FBS, which was guaranteed doxycycline-free (PAN Biotech GmbH, Germany). MESCs were cultured in high glucose DMEM medium supplemented with 4 mM L-Glutamine (PAA M11-006), 1x Penicillin/Streptomycin, 1 mM sodium pyruvate solution, 1x non-essential amino acids (ThermoFisher Scientific, USA), 100 µM sterile-filtered β-mercaptoethanol (in water), 0.41% v/v LIF (self-made, supernatant from LIF-producing cells, batch: 26.7.14 from Dennis Sadic, Schotta lab, LMU Munich) and 15% FBS (Sigma F7524, lot: 091M3398) in 0.2% gelatin/PBS-coated dishes. To passage, collect or seed cells, they were washed briefly with phosphate-buffered saline (PBS) and detached from the culture dish by incubation with 0.05% Trypsin/0.02% ethylenediaminetetraacetic acid (EDTA) in PBS (PAN Biotech GmbH, Germany) for 2-5 min at 37°C. For immunofluorescence and/or RNA fluorescence in situ hybridization (FISH) staining, cells were seeded on 12 mm diameter glass slides. For live cell experiments, they were seeded either in 8-well Lab-Tek chambers (#154534, ThermoFisher Scientific Inc., USA) or 10-well CELLview slides (#543079, Greiner Bio-One GmbH, Germany) that were coated with 1% Matrigel in serum-free DMEM beforehand to allow for better cell attachment and limit cell movement during image acquisition. U2OS 2-6-3 cells were transfected with Effectene (Qiagen, Netherlands) whereas NIH3T3 and iMEFs were transfected with Lipofectamine (LTX or 3000, ThermoFisher Scientific, USA) according to the manufacturer's protocols. For qRT-PCR of transfected U2OS 2-6-3 cells, Xtreme gene 9 (Roche, Germany) was used for transfection. For some experiments, cells were pre-treated with the histone deacetylase (HDAC) inhibitor suberanilohydroxamic acid (SAHA). In that case, fresh medium containing 2 µM SAHA (Millipore) in ethanol was added 4-5 h after transfection (if applicable) and ∼24 h before the respective experiment (microscopy, western blot or immunofluorescence staining).

IV.2.3 Cell fixation and immunofluorescence staining

Cells on coverslips were briefly washed with PBS, incubated with 4% paraformaldehyde (PFA) in PBS (Santa Cruz, Germany) for 7-12 min at room temperature (RT) and washed again with PBS (three times). For immunofluorescence staining, cells were permeabilized with 0.1% TritonX-100 (Merck, Germany) in PBS for 5 min and washed twice with PBS. Fixed cells were then incubated 30-60 min with 10% goat serum (Sigma-Aldrich, USA) in PBS to block unspecific antibody binding sites. The appropriate primary antibodies (**Table 9**) were diluted as indicated in 5% goat serum. Incubation was either 60-120 min at RT or over night at 4°C. After washing thrice with PBS/0.002% NP-40 (Sigma-Aldrich, USA), the cells were incubated with fluorophore-coupled secondary antibodies (**Table 9**) diluted in PBS for 45-90 min at RT. After washing another three times with PBS and rinsing the coverslips briefly with water to desalt and 70% followed by 100% ethanol to

dry, cells were mounted on a microscopy slide with a drop of Prolong Gold Antifade mountant containing DAPI (ThermoFisher Scientific Inc., USA).

IV.2.4 RNA fluorescence *in situ* hybridization

RNA FISH was performed on cells seeded (and transfected/treated) on glass coverslips. After washing the cells once with PBS, they were incubated with CSK buffer with 10 mM vanadyl ribonucleoside complex (VRC) for RNase inhibition or 50 µg/ml RNase A (negative control, ThermoFisher Scientific, USA) on ice for 5 min. Another washing with PBS was followed by fixation with 4% PFA for 12 min. The cells were washed again with PBS and dried with an ethanol series (70% → 85% → 100%; 3 min each). Meanwhile, 50 ng of the Atto565-labeled MS2 probe (**Table 11**) were mixed with 10 µg salmon sperm DNA and 5 µl formamide (all quantities are per slide) and incubated first at 37°C for 10 min and then at 74°C for 7 min. Last, the probe mix was diluted 1:1 with 2× hybridization buffer and supplemented with 10 mM VRC or 50 µg/ml RNase A (ThermoFisher Scientific, USA). Cells were incubated with the probes in a wet chamber at 37°C over night. On the next day, they were washed twice with 2× SSC buffer/50% v/v formamide at RT for 15 min, once with 0.2× SSC/0.1% v/v Tween-20 at 40°C for 10 min, twice with 2× SSC at RT for 5 min and once with PBS. This protocol destroyed fluorescent proteins, so that it was generally followed by an immunofluorescence staining using a specific antibody against GFP (to detect GFP and/or YFP). Therefore, cells were permeabilized with ice-cold 0.1% v/v Triton X100/PBS for 5 min and then the immunofluorescence protocol described above was followed.

IV.2.5 Confocal fluorescence microscopy

All fluorescence microscopy data was acquired using microscopes from the DKFZ light microscopy core facility. FRAP and BLInCR experiments were carried out on a Leica SP5 microscope equipped with a HCX PL APO lambda blue 63.0×1.40 OIL UV objective, a UV diode (405 nm), a multiline argon laser (458 nm, 476 nm, 488 nm, 496 nm, 514 nm) and three HeNe lasers for excitation at 561 nm, 594 nm and 633 nm. This setup was also used for imaging except for H3K9me3/HP1 imaging in different iMEF cell lines and Pax imaging in NIH3T3 cells. Those experiments were done using the Zeiss LSM 710 microscope equipped with a 63x/1.4 NA oil DIC III objective and an avalanche photodiode (APD). A summary of the microscope settings used for the respective experiments is given in **Table 12**. Fluorescent proteins and dyes with their excitation and emission wavelengths used in this study are listed in **Table 13**. When imaging more than one fluorescent protein, imaging was done sequentially.

Table 12: Microscopy settings for different acquisitions.

experiment	microscope	image size (px)	scan speed	line average	z stacks
Imaging	Leica SP5	512×512	400 Hz	4×	Δz=0.5 µm
APD imaging	Zeiss LSM710	512×512	speed 3 or 5	4×	no
FRAP	Leica SP5	128×128	1400 Hz	1×	no
BLInCR					
recruitment		256×256	1400 Hz	1×	no
reversibility	Leica SP5	512×512	400 Hz	2×	Δz=0.5 µm, #z=7
txn activation		512×512	400 Hz	4×	Δz=0.5 µm, #z=4-5

More specific parameters are listed in the respective sections.

Table 13: Excitation & emission settings for imaging fluorescent proteins and dyes.

fluorescent protein/dye	excitation wavelength (nm)	detected emission range (nm)
tagBFP, CFP, DAPI	405	415-475
GFP, Alexa488	488	500-550
YFP	488 (or 514)	500-550/75 (or 525-575)
tagRFP, tagRFP-T, Alexa568	561	575-630
mCherry, mKate2 *)	561 or 594	575/605-750 or 570-615 **)
iRFP713, Alexa647	633	645-780

*) Excitation at 561 nm resulted in a better signal (higher signal-to-noise ratio), but also caused more photobleaching. Therefore, mKate2 was excited at 594 nm for time lapse experiments. **) When the sample was stained with an Alexa647-linked antibody also, this emission range was used to avoid bleed-through as Alexa647 can also be excited at 561 nm.

IV.2.6 PCH enrichment analysis

To quantify the enrichment of factors at PCH from CLSM images, the DAPI image was used for threshold-based selection of PCH using ImageJ (Schneider *et al.* 2012). The entire nucleus apart from the PCH foci was considered *bona fide* euchromatin (**Fig. 42**). The resulting masks were then used to measure the intensity of the factor of interest (e. g. H3K9me3). The PCH enrichment d of any factor was calculated from the mean PCH and euchromatin intensities I_{PCH} and I_{Eu}: $d = I_{\mathrm{PCH}}/I_{\mathrm{Eu}}$. This enrichment was normalized to the chromatin density c which can be calculated from the measured DAPI enrichment $d(\mathrm{DAPI})$ as $c = d_{\mathrm{ref}}(\mathrm{H2A}) \cdot d(\mathrm{DAPI})/d_{\mathrm{ref}}(\mathrm{DAPI})$. Here, the reference H2A and DAPI enrichments were $d_{\mathrm{ref}}(\mathrm{H2A}) = 1.80.1$ and $d_{\mathrm{ref}}(\mathrm{DAPI}) = 2.7 \pm 0.1$. Those values were obtained from imaging DAPI and mRFP in an NIH3T3 cell line stably expressing H2A-mRFP (Muller-Ott *et al.* 2014). The corrected enrichment was thus calculated as $d_{\mathrm{corr}} = d/c$.

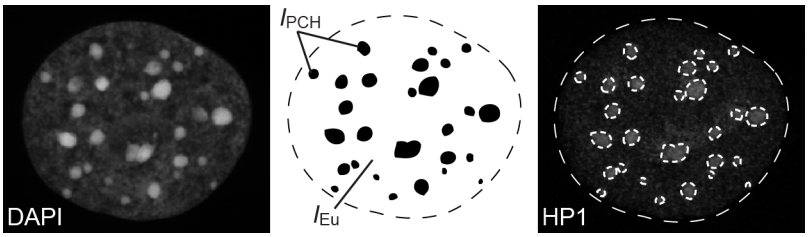

Figure 42: Quantification of PCH enrichment. PCH foci were identified in the DAPI image (left) by using a threshold-based approach. The detected foci (center) were despeckled (`Process` → `Noise` → `Despeckle`) and used as a mask to quantify the fluorescence intensity of a factor of interest (right). Figure adapted from (Muller-Ott *et al.* 2014).

IV.2.7 Fluorescence recovery after photobleaching

For fluorescence recovery after photobleaching (FRAP), fluorescent proteins in a circular region of interest (ROI) within the nucleus were bleached and recovery of fluorescence intensity at the bleach area was recorded as depicted for an exemplary cell in **Figure 43** (top). From the recovery curve (**Fig. 43**, bottom), protein mobility parameters such as the (effective) diffusion coefficient and chromatin binding/dissociation rates can be derived ((Mueller *et al.* 2010, Vivante *et al.* 2017)).

Data acquisition. Cells were transfected with the GFP-tagged protein of interest (i. e. Pax3-GFP or GFP-MeCP2) and the FRAP Wizard included in the Leica imaging software (LAS AF) was used with the general parameters listed in **Table 12** resulting in a frame time of 0.115 s. Photobleaching was done in an area of 1.5 µm in diameter either at PCH foci or in euchromatic regions. The Argon laser was set to 80% in the configuration and the laser lines at 458 nm, 476 nm, 488 nm and 496 nm were used at 100% in the acquisition setup for two bleaching frames. Pre- and post-bleach imaging was done using an appropriate laser power depending on the expression level of the fluorescently tagged protein. The number of pre- and post-bleach frames was 50 and 700, respectively, for Pax3-GFP. For GFP-MeCP2, 70 pre-bleach frames were acquired and the number of post-bleach frames was 600 for euchromatin and 1200 for PCH.

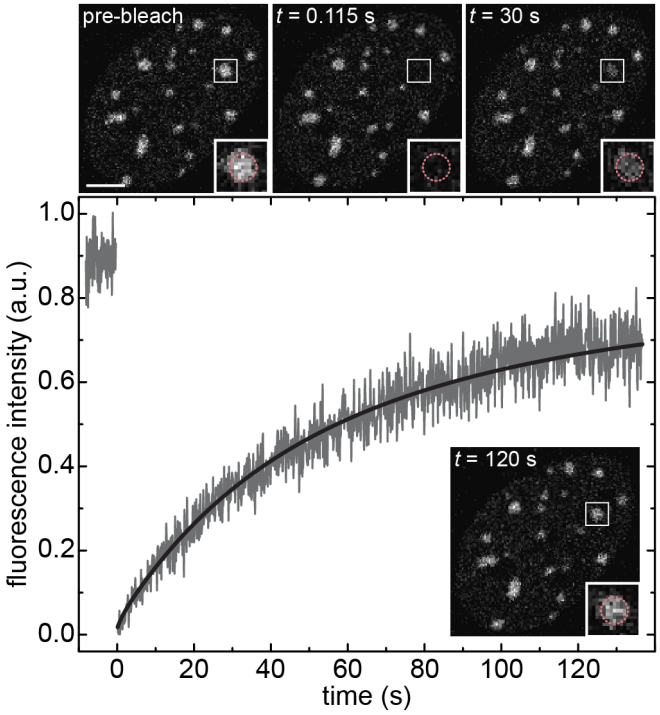

Figure 43: Principle of FRAP. Fluorescent proteins in the ROI (dashed pink) were photobleached using high laser intensity. The recovery of fluorescence intensity at the ROI was recorded and the FRAP curve was fitted using FREDIS ((Muller *et al.* 2009)). In this example, GFP-MeCP2 at a PCH focus was bleached in a wt iMEF. Scale bar: 5 µm, gray: experimental data, black: reaction-diffusion fit.

Image analysis. To account for cell movement during the acquisition, images were registered using the TurboReg plugin ((Thevenaz *et al.* 1998)) in the Fiji distribution ((Schindelin *et al.* 2012)) of ImageJ ((Schneider *et al.* 2012)). The maximum projection of the pre-bleach frames was used as target. In addition, the bleach profile was analyzed since it is taken into account in a newer version of the FREDIS software ((Muller *et al.* 2009)). Due to surrounding heterochromatin foci, bleach profile analysis based on the first post-bleach frame as suggested by Mueller and colleagues ((Mueller *et al.* 2008)) was not feasible for PCH FRAP. Therefore, the first post-bleach image was subtracted from the last pre-bleach image using ImageJ (`Process` → `Image Calculator...`) and the radial profile was calculated from this difference image (**Fig. 44a**) with a radius of 15 px (`Plugins` → `Radial Profile`). For euchromatin FRAP, this analysis was done on the first post-bleach image. All profiles from one condition (i. e. PCH or euchromatin in the same cell type) were normalized to the average intensity of a distant region (**Fig. 44a**), averaged (**Fig. 44b**) and fitted with the following constant function with Gaussian edges based on ((Mueller *et al.* 2008)) using `R`:

$$I(r) = \begin{cases} b & \text{for } r \leq r_c \\ \theta \left(1 - \exp\left(\frac{-(r-r_c)^2}{2\sigma^2}\right)\right) & \text{for } r > r_c \end{cases} \quad (1)$$

Here, $\theta - b$ is the bleach depth, r_c the radius of the constant and σ the width of the Gaussian. The latter two parameters were used as input parameters for FREDIS and are listed in **Table 14**. Note that the average profile was subtracted from 1 to yield the curve as in **Figure 44b** when the image difference was used (i. e. for PCH FRAP).

Table 14: Bleach profile parameters.

construct	sample		σ (µm)	r_c (µm)
GFP-MeCP2	PCH	wt	0.39	0.15
		dn	0.38	0.16
	euchromatin	wt	0.59	0.18
		dn	0.88	0
Pax3-GFP	PCH		0.80	0.07
	euchromatin		1.01	0

The profile analysis was done using either the first post-bleach image (euchromatin) or the difference of the last pre- and the first post-bleach image (PCH). The parameters were subsequently used for FRAP analysis with FREDIS.

Data analysis. Registered time series were analyzed using the FREDIS software ((Muller *et al.* 2009)) that is based on the theoretical framework previously described ((Sprague *et al.* 2004)). Based on more recent advances in FRAP data analysis ((Mueller *et al.* 2008)), the newest FREDIS version used here also takes the shape of the initial bleach

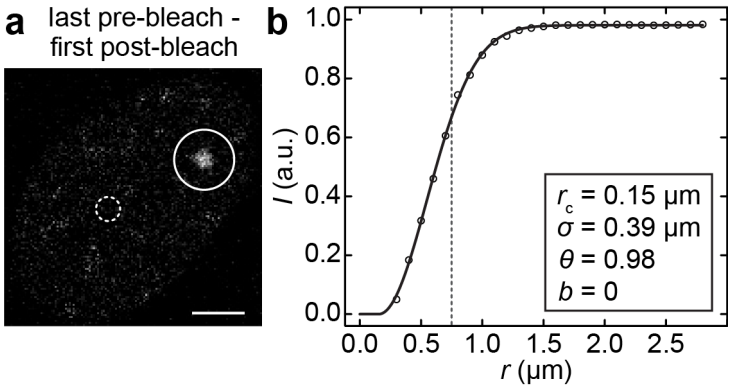

Figure 44: Bleach profile analysis for FRAP. (**a**) First post-bleach image subtracted from last pre-bleach image (same cell as in **Fig. 43**) to analyze the average radial profile (solid circle) using ImageJ. The mean intensity in the reference area (dashed circle) was used for normalization. This example shows GFP-MeCP2 FRAP at a PCH focus. Scale bar: 5 µm. (**b**) Average of the GFP-MeCP2 PCH FRAP profiles in wt iMEFs. The curve was fitted with a constant function with Gaussian edges yielding the radius of the constant r_c and the width of the Gaussian σ that were used as input parameters for subsequent fitting of FRAP curves using FREDIS. The dashed line represents the actual bleach radius (0.75 µm).

profile into account. All curves were fitted with the three models implemented in FREDIS, namely a binding model, a diffusion model and a combined reaction-diffusion model. To determine if the more complex model (i. e. the reaction-diffusion model) yielded a significantly better fit, a ratio F and the resulting P value were calculated as in ((Muller *et al.* 2009)):

$$F = \frac{\chi^2_s - \chi^2_{rd}}{\chi^2_{rd}} \cdot \frac{f_{rd}}{f_s - f_{rd}} = \frac{\chi^2_s - \chi^2_{rd}}{\chi^2_{rd}} \cdot \frac{n - p_{rd}}{p_{rd} - p_s} \tag{2}$$

$$P = 1 - cdf(F, f_s - f_{rd}, f_{rd}) \tag{3}$$

Here, χ^2 was the sum of the residuals as calculated by FREDIS for the simple model (s, i. e. the reaction or diffusion model) and the complex model (rd, i. e. the reaction-diffusion model), respectively. The degrees of freedom f can be expressed in terms of the number of data points n and the number of fit parameters p ($p_s = 1$, $p_{rd} = 3$). To calculate the P value, the cumulative distribution function cdf (implemented in `R` as `pf()`) was used. If $P \leq 0.05$, the reaction-diffusion model was considered to fit the data significantly better than the respective simple model. To calculated average curves for the FRAP experiments, the FREDIS output curves (**Fig. 43**) were normalized to the average pre-bleach intensity. The mean of the thus normalized curves was refitted using FREDIS and the best fit is displayed in all average curves (**Fig. 10c** & **Fig. 11b**).

IV.2.8 Blue Light-Induced Chromatin Recruitment (BLInCR)

BLInCR is a tool devised in this thesis to rapidly increase the local concentration of effector proteins at subnuclear loci upon illumination with blue light (Rademacher *et al.*

2017).

Data acquisition. U2OS 2-6-3 cells or MEF cell lines (wt iMEFs or NIH3T3) were transfected with the appropriate BLInCR constructs from **Table 6** and **Table 7** and treated with the histone deacetylase inhibitor SAHA if needed. To avoid pre-recruitment of PHR-tagged effectors, light exposure of transfected cells had to be avoided. In particular, cells were kept in a dark box after transfection and mounted on the microscope using a red flashlight. BLInCR images were recorded with the parameters given in **Table 12** and **Table 13**. Generally, a zoom factor of 9 corresponding to 53.39 nm/px was used. Excitation of fluorophores with wavelengths ≥561 nm did not cause PHR switching and was thus suited to follow effector and/or reader dissociation from the array. For longer time series experiments (BLInCR dissociation kinetics and transcription activation), z stacks were recorded and the focus was adjusted if necessary.

Image analysis. Image processing and intensity quantifications were done using the Fiji distribution (Schindelin *et al.* 2012) of ImageJ (Schneider *et al.* 2012). When z stacks were recorded, they were registered using the StackReg plugin for single channel stacks (reversibility kinetics) or the TurboReg plugin for multi channel stacks (Thevenaz *et al.* 1998). For intensity quantifications, a maximum projection of the z stacks was used (`Image` → `Stacks` → `Z Project...`). The mean fluorescence intensity was measured (`Analyze` → `Measure`) at a nuclear reference region (30/60 px for 256×256/512×512 px images, respectively) and the array (15/20-40 px for 256×256/512×512 px images, respectively). The nuclear reference region was selected to be in a homogeneously stained region (e. g. not at nucleoli) and close to the array to minimize possible intensity differences due to inhomogeneous illumination of the cell. For time-lapse experiments, the enrichment (of the effector or reader) at the array $E(t)$ was calculated as the difference of the mean fluorescence intensities at the array region I_{array} and the reference region I_{ref}. The mean intensity at the nuclear reference region I_{ref} and at a background region I_{back} outside of the cell were used to account for bleaching by fitting their difference to an exponential decay ($I_{\text{ref}} - I_{\text{back}} = a \cdot e^{-k \cdot t}$). The bleach-corrected enrichment was thus calculated as $E(t) = (I_{\text{array}} - I_{\text{ref}})/e^{-k \cdot t}$.

BLInCR kinetics. The array was marked with CIBN-TetR-tagRFP-T and could thus be localized without triggering PHR switching as depicted in **Figure 25b** (Results section). Subsequently, a YFP time series of 400 frames with 204.3 ms frame time was acquired (see parameters in **Table 12** and **Table 13**) thereby triggering and recording PHR-YFP-effector switching and recruitment. The enrichment of PHR-YFP-labeled effectors at the array $E(t)$ at time t after the first blue light exposure was calculated as described above, but without bleach correction since the measurement time was rather short and YFP rather bleach-resistant (see Results section, **Fig. 25b**). A maximum projection of the time series was used to select array and reference region. The enrichment at the array E after time t ($E(t) = (I_{\text{array}} - I_{\text{ref}})$ was fitted to a model considering two

parallel first-order reactions:

$$E(t) = a - b \cdot e^{k_1 \cdot t} - c \cdot e^{k_2 \cdot t} \tag{4}$$

Each single curve was fitted using a non-linear generalized least squares method (gnls) implemented in the `nlme` package in `R` (Pinheiro *et al.* 2017). In an alternative approach, the two characteristic rates k_1 and k_2 were kept constant for all cells and all constructs under the assumption that they correspond to the PHR-CIBN heterodimerization and the PHR oligomerization and are therefore identical for all effector constructs. Using the global fitter (by F. Erdel), all 123 single recruitment curves were fitted with individual constants a, b and c to account for expression and effector differences, but with global rates k_1 and k_2. To assess the goodness of fit, the cumulative sum of squared residuals of the global fit were devided by the cumulative sum of squared residuals of the single fits. In addition, the coefficient of determination R^2 was calculated for each single curve and the distribution of those coefficients is shown in **Figure 25f**. For both fit strategies, the characteristic time to reach half maximal levels $\tau_{1/2}$ was calculated from the model (**Eq. 4**) and the half-maximal level $E(\tau_{1/2})$:

$$E(\tau_{1/2}) = E(0) + \frac{E(\infty) - E(0)}{2} \tag{5}$$

With the offset $E(0) = a - b - c$ and the plateau value $E(\infty) = a$, $\tau_{1/2}$ was determined using the `uniroot` function in `R` to solve

$$0 = b \cdot (1 - 2e^{-k_1 \cdot \tau_{1/2}}) + c \cdot (1 - 2e^{-k_2 \cdot \tau_{1/2}}) \tag{6}$$

Single curves were normalized to the plateau value a ($E_{\mathrm{norm}}(t) = E(t)/a$) to account for different transfection efficiencies before averaging. Cells that moved during image acquisition or had a very low signal-to-noise ratio were removed from the analysis.

BLInCR dissociation kinetics. To characterize the BLInCR dissociation kinetics, PHR-mCherry-effector constructs were used to avoid triggering re-recruitment during imaging of those constructs. They were co-transfected with CIBN-TetR ('localizer') and an GFP-LacI (array marker) and imaged with the appropriate parameters (**Table 12** & **Table 13**). A pre recruitment z stack was recorded before triggering relocalization of PHR-mCherry-tagged effectors by imaging the array marker (excitation at 488 nm, two sequential stacks with a total of 38 s exposure to blue light). The first post recruitment stack in the mCherry channel was recorded right after the two recruitment stacks and constitutes the time point $t = 0$. Subsequent stacks were recorded every ∼30 s for the first 5 min and then at longer time intervals for a total of ∼30 min. Image analysis was done as described above. The array and reference region were selected manually for each time point since the cell shape changed considerably during the acquisition period

rendering registration of the time series impractical. Importantly, the size of the array measurement area was large enough to include the entire array and was always the same for a single cell time trace. The bleach-corrected enrichment $E(t)$ of single cells was fitted with a model considering a single exponential with a time-dependent (i. e. concentration-dependent) rate similar to a previously proposed model (Sing *et al.* 2014):

$$E(t) = a \cdot e^{-k \cdot t^m} + c \tag{7}$$

The characteristic half-life time $t_{1/2}$ was calculated from the model fit (**Eq. 7**) and the half-maximal level $E(t_{1/2})$ (as **Eq. 5**). With the initial enrichment at $E(0) = a + c$ and the offset $E(\infty) = c$, $t_{1/2}$ was determined using the `uniroot` function in `R` to solve

$$0 = a \cdot \left(e^{-k \cdot t_{1/2}^m} - \frac{1}{2} \right) \tag{8}$$

Cells were normalized to account for different transfection efficiencies ($E_{norm}(t) = (E(t) - c)/a$) and points at regular intervals were interpolated (linear interpolation between neighboring points) to allow for averaging of single traces.

Transcription activation with VP16 BLInCR. Cells were transfected with the 'localizer' CIBN-TetR, the effector construct PHR-YFP-VP16, the array marker CFP-LacI and the RNA reader domain mKate2-MCP. A z stack of the latter construct was recorded before the initial recruitment of the effector by imaging of YFP and CFP ("between stacks" sequential scan mode). Subsequently, stacks in all three channels ("between lines" sequential scan mode) were recorded every 2-4 min thereby consistently re-recruiting VP16 to the reporter array. Reader enrichment was quantified as described above and normalized to the level before the first recruitment and after 50 min. For normalization and averaging, time points every 2.5 min were interpolated linearly between neighboring time points. The averaged time traces were fitted to three different models: a two-state model, a model including positive feedback and a sequential activation model (fitting by Fabian Erdel, see (Rademacher et al. 2017) for details). The inflection points (**Fig. 36c**) were calculated from the feedback model fit (by Fabian Erdel). To determine the array size (**Fig. 36b**), a local area around the array was selected (60 px/3.2 µm in diameter) either in the YFP or CFP channel and converted to a binary image using Otsu's method (Otsu 1979) for thresholding (`Image` → `Adjust` → `Threshold...`). The size was measured (`Analyze` → `Measure`) and normalized to the value at 2.5 min to ensure that recruitment was complete at the first measurement time point (which was critical when using the PHR-YFP-VP16 images).

Constitutive BLInCR with LED illumination over night. Control experiments for VP16 BLInCR were conducted with the same constructs as the time-resolved VP16 BLInCR, but the cells were illuminated over night with a blue LED. To test reader do-

mains such as fluorescently tagged hADD or mCD (see Results section and **Table 6** for details), cells were seeded, transfected and imaged in 8-well chambers suited for CLSM (Labtek, ThermoFisher Scientific, USA). Cells that were subsequentially stained for histone PTMs or pol II phosphorylation were seeded on 12 mm glass coverslips and treated for IF as described above. Cells were seeded in a removable 4-well chamber (ibidi, Germany) for direct comparison of H3K9ac levels by IF (**Fig. 35**). The chamber was removed after washing once with PBS before adding PFA. After over night activation, fixation and IF (when applicable), images were recorded as specified in **Table 12**, processed and the mean fluorescence intensity at a nuclear reference area I_{ref} was measured as described above. To analyze the enrichment of the reader or histone modifications compared to the local chromatin environment, radial profiles around the array were calculated (diameter: 70 px/3.7 µm) yielding the intensity $I(r)$ at distance r from the array center which was normalized to I_{ref} ($I_{\text{norm}}(r) = I(r)/I_{\text{ref}}$). For fixed samples the reader profile was devided by the normalized DAPI profile to account for chromatin density. To quantify the accumulation of the reader proteins at the array, the enrichment over the reference $I_{\text{norm}}(r)$ was integrated from the array center to the FWHM yielding the accumulated enrichment E_{sum}:

$$E_{\text{sum}} = \sum_{r=0}^{r=FWHM} 2\pi r I_{\text{norm}} \tag{9}$$

The FWHM was calculated from the effector protein profile to account for different array sizes resulting from recruitment of different effectors.

IV.2.9 Quantitative real-time polymerase chain reaction (qRT-PCR)

Experimental procedure. Total RNA was isolated from $0.5 - 1.5 \times 10^6$ cells using the Nucleospin RNA plus kit (Macherey Nagel, Germany) according to the manufacturer's protocols. Alternatively, long RNA (>200 nt) isolated for RNA-seq (see below) was used. Complementary DNA (cDNA) was synthesized from 650 ng – 1 µg RNA using the SuperScript III or IV reverse transcriptases (ThermoFisher Scientific, USA) with random hexamer primers according to the manufacturer's protocols. Samples without reverse transcriptase were prepared as controls. For the calibration standard, 2-2.5 µg RNA from SUV39H dn iMEFs was used. After digestion of the RNA templates with 2.5 U RNase H (New England Biolabs, USA), the standard was diluted sequentially (1:3), resulting in a total of five to eight standards. qRT-PCR was performed using the SYBR green PCR master mix (Roche, Germany or Applied Biosystems, USA) and the primers for GSATs, Actin, GAPDH and the CFP-SKL reporter gene in the U2OS 2-6-3 cell line are listed in **Table 11** with the following PCR program on a step one plus real-time PCR light cycler (Applied Biosystems, USA): 95°C/10 min (step 1), 95°C/10 s (step 2), 60°C/10 s (step 3), 72°C/10 s (step 4) with cycling from step 4 to step 2 for a total of 28 (GSAT) or 35 (Actin, GAPDH) cycles. All samples and the standard were analyzed in triplicates

and the controls without reverse transcriptase were analyzed in duplicates or as single controls.

Data analysis. The raw data analysis was done with the step one software (Applied Biosystems, USA) and the normalization was done in Excel. Briefly, the standard curves were used to determine quantities from the measured C_T values. Mean GSAT or CFP-SKL quantities from the three technical replicates were normalized to the respective actin or GAPDH quantities. The fold change of those normalized quantities compared to the average of the biological replicates of wild type iMEFs (**Fig. 18a**), untransfected cells (**Fig. 27f**, **Fig. 28a**) or non-activated values ("pre" in **Fig. 37**) are shown, respectively. The error bars reflect the sem of the biological replicates.

IV.2.10 Histone extraction and western blot

Acid extraction of histones. For western blot protein analysis of modified histones, the latter were isolated by an acid extraction protocol based on published protocols ((Shechter *et al.* 2007, Villar-Garea *et al.* 2008)). Briefly, 10^6-10^7 cells were washed with PBS and resuspended in 0.25 M HCl (100 µl/10^6 cells) and rotated over night at 4°C. After centrifugation at $20,800 \times g$ at 4°C in a table-top centrifuge for 7-10 min, the supernatant was set aside (4°C or on ice) and the pellet was redissolved in the half the volume HCl and rotated at 4°C for another 3-4 h. The supernatants were combined and histones were precipitated for ∼2 h on ice (or at 4°C) by adding trichloroacetic acid (TCA) to a final concentration of 33% v/v. After centrifuging for 20-60 min at 4°C and 13000 rpm, the supernatant was discarded and the pellet washed twice with 170 µl and 150 µl *tert*-butyl methyl ether (TBME)/ethanol (1:1). The pellet was dried (∼20 min at RT), dissolved in 1× Laemmli buffer (Alfa Aesar, USA) to a final "cell concentration" of 10^4 cells/µl and heated to 100°C for 10 min.

Protein separation and western blot. Acid-extracted proteins were separated on a stainfree gel (4-20%, Bio-Rad Laboratories Inc., USA) for 45 min at 160 V and semi-dryly blotted onto a nitrocellulose membrane (Bio-Rad Laboratories Inc., USA) for 1 h at 200 mA. After transfer, unspecific binding sites were blocked with 5% milk in TBS for 1 h at RT. The primary antibodies against H3K9me3 (see **Table 9**) were diluted in 1% bovine serum albumin (BSA) in TBS with 0.02% sodium azide and incubated with the membrane at 4°C over night. After washing thrice with 0.1% Tween in TBS, the membrane was incubated with a secondary antibody coupled to HRP in 5% milk/TBS for 1 h at RT. After washing again, the secondary antibody was detected by chemoluminescence (Clarity Western ECL, #170-5060, Bio-Rad Laboratories Inc., USA) using a ChemiDoc MP Imaging System (Bio-Rad Laboratories Inc., USA). To check protein loading, antibodies were stripped from the membrane (30 min at 56°C, Roti-free stripping buffer, Carl Roth GmbH & Co. KG, Germany). Subsequently, the membrane was blocked again and incubated with a primary antibody against H3 following the same protocol as described

above.

Histone tail modification antibody tests on the modified peptide array. The specificity of primary antibodies against H3K9me3 (**Table 9**) was tested on a modified histone peptide array (#13005, ActiveMotif) by following the same procedure as described in the previous section for protein detection on the membrane after western blotting. Notably, antibodies could also be stripped from the peptide array so that the latter could be reused at least three times.

IV.2.11 Chromatin immunoprecipitation followed by DNA sequencing

ChIP. $3 - 10 \times 10^6$ cells were pelleted, dissolved in 1% formaldehyde (#28906, ThermoFisher Scientific Inc., USA) in PBS and incubated on a rotating mixer for 10 min at RT. The fixation reaction was stopped by addition of glycine to a final concentration of 125 mM. After washing the cells twice with PBS/0.5 mM phenylmethylsulfonyl fluoride (PMSF), they were dissolved in swelling buffer (**Table 10**) and incubated on ice for 10 min. The cells were pelleted, dissolved in MNase buffer (**Table 10**) and incubated with 4 U MNase (ThermoFisher Scientific, USA) at 37°C for 15 min. Samples were stored on ice and diluted with the appropriate amount of 10× sonication buffer (to yield 1×**Table 10**). The samples were transferred to glass tubes (Covaris Inc., USA) and sonicated in a Covaris S2 sonicator (parameters: 900¸s, burst 200, cycle 20%, intensity 8). After centrifugation at 13000 rpm and 4°C for 15 min to pellet cell debris, the supernatant was used for immunoprecipitation. ChIP-grade protein G magnetic beads (25 µl/sample, #9006S, Cell Signalling, UK) were washed twice and dissolved in sonication buffer with 1× protease inhibitor. The cell supernatant and 4 µg normal rabbit IgG (lot: ER1212071) were added to the beads and the mixture was rotated at 4°C for ∼2 h (pre-clearance). About 1/20 of the supernatant was used as input sample and the remainder was split equally for three ChIP reactions using 4 µg of the respective antibody (IgG, H3 and H3K9me3) and incubating at 4°C for 2 h while rotating. Meanwhile, protein G magnetic beads were prepared as described above and then added to each IP sample. The precleared sample, antibody and bead mixture was incubated on a rotator at 4°C over night. Meanwhile, the input samples were filled up to 500 µl with elution buffer (**Fig. 10**) and RNaseA (ThermoFisher Scientific, USA) was added to a final concentration of 10 µg/ml. After one hour at 37°C, Proteinase K (M3037, Genaxxon bioscience, Germany) and NaCl were added to final concentrations of 80 µg/ml and 200 mM, respectively and the input samples were incubated over night at 65°C (protein digestion and reverse crosslink). On the next day, the IPs were washed once with sonication buffer, once with high-salt buffer, once with Li buffer and twice with TE buffer (see **Table 10** for buffer compositions). For each washing step, the samples were rotated and incubated for 5 min at RT and then placed on a magnetic rack. Subsequently, the DNA-antibody complexes were eluted twice with elution buffer (50 mM Tris pH8.0, 1 mM EDTA, 1% SDS, 50 mM $NaHCO_3$)

at 37°C (15 min) on a shaker to avoid bead sedimentation. The IP samples were treated with RNaseA and Proteinase K as described above for the input samples. Meanwhile, the input DNA was precipitated by adding 0.1 volume 7.5 M ammonium acetate, 1 volume isopropanol and 1.5 µl Glycoblue and storing the samples at -20°C for ∼75 min. The precipitated DNA was pelleted by centrifugation at 4°C and full speed for 40 min and the pellet was washed once with ice-cold 75% Ethanol before dissolving it in water. The input samples were used to assess the fragment size by gel electrophoresis on a 2% agarose E-Gel (ThermoFisher Scientific, USA) and on a DNA1000 bioanalyzer chip (Agilent Technologies Inc., USA). On the next day, the IP samples were precipitated as the input samples on the previous day and their concentrations were measured by Qubit (dsDNA or dsDNA high sensitivity, ThermoFisher Scientific, USA).

Library preparation. The libraries were prepared using ∼50-800 ng of precipitated DNA (or all if the concentration could not be measured by Qubit (dsDNA or dsDNA high sensitivity, ThermoFisher Scientific, USA)) using the NEBNext Ultra DNA Library Prep Kit for Illumina (#E7370, New England Biolabs Inc., USA) following the manufacturers' protocols. Notably, diluted adapters (1.5 µM) were used for the adapter ligation and insert fragments of 150 bp (i. e. mononucleosomes) corresponding to 270 bp total library size (including insert and adaptors) were selected in the size selection step. For library amplification, 13 PCR cycles were carried out. Lastly, the library size and quality were checked on DNA1000 bioanalyzer chips. Libraries were sequenced and demultiplexed by the DKFZ core facility on a HiSeq 2000 machine (Illumina, USA).

Initial data processing. All samples were mapped with Bowtie ((Langmead *et al.* 2009)) according to the mapping scheme depicted in **Figure 45**. Briefly, reads were mapped to the mouse major satellite consensus sequence ((Lehnertz *et al.* 2003)) without mismatches. All unmapped reads were then mapped to the mm9 assembly of the reference genome with random assignment of multiple mapping reads and without mismatches. Lastly, unmapped reads were remapped allowing up to two mismatches. The raw read numbers and resulting read distributions are shown below (**Fig. 46**) and are listed in the Appendix (**Supplementary Table S2**). Files containing mapped reads (with or without mismatches) were sorted by genomic coordinate, merged and indexed using samtools ((Li *et al.* 2009)). For quality control, the Pearson correlation of read counts within 10 kb windows across the entire genome were calculated using the `multiBamSummary` and `plotCorrelation` functions of deepTools2 (Ramirez *et al.* 2016). As another quality control measure, fingerprints were calculated using the `plotFingerprint` function of deepTools2.

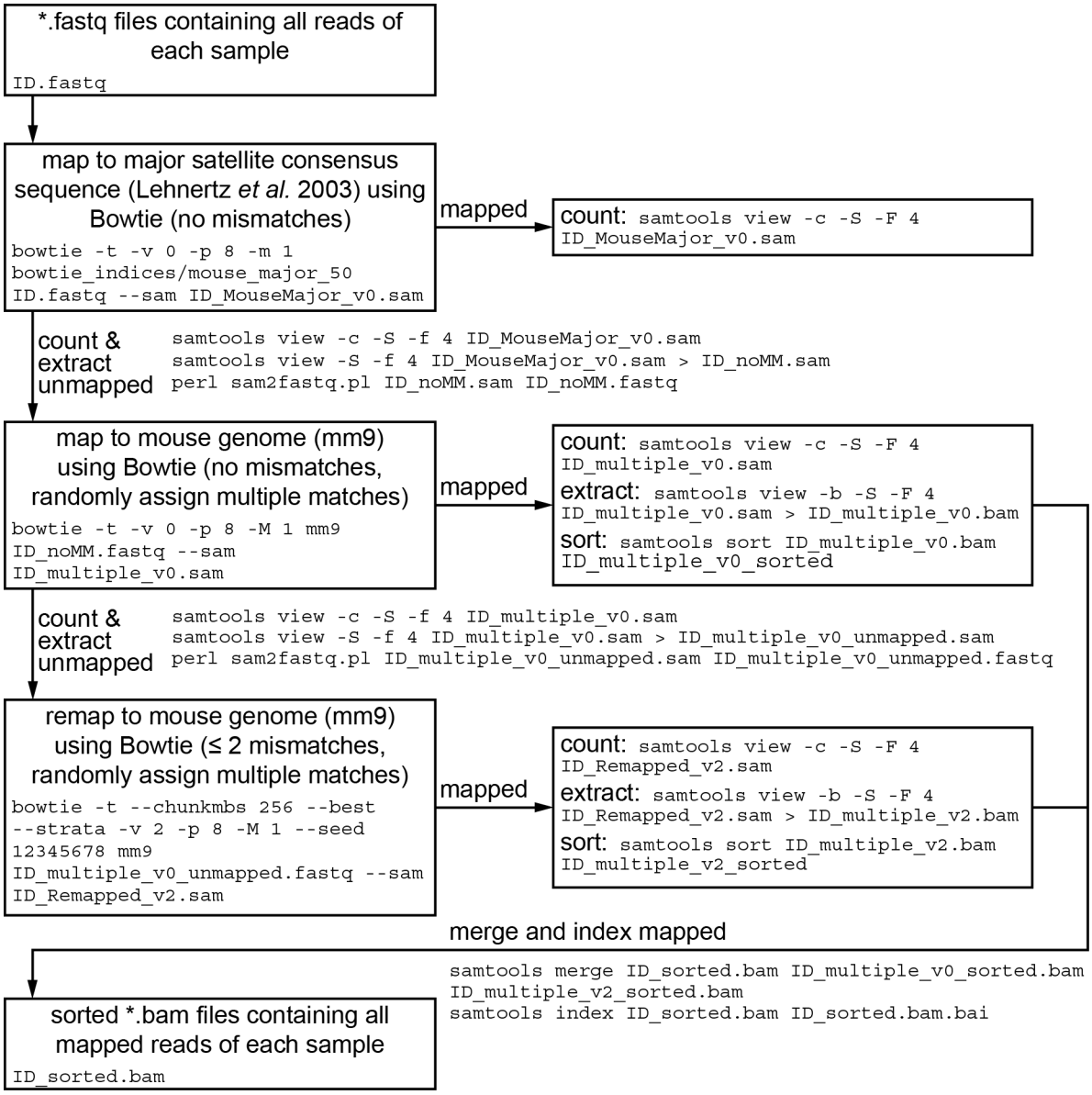

Figure 45: Mapping strategy for reads from ChIP-seq. All samples (i. e. input, IgG control, H3 ChIP and H3K9me3 ChIP of all cell lines) were mapped according to the scheme depicted here using Bowtie. An overview of the mapped reads (unique/multiple, with/without mismatches) is given in **Fig 46**.

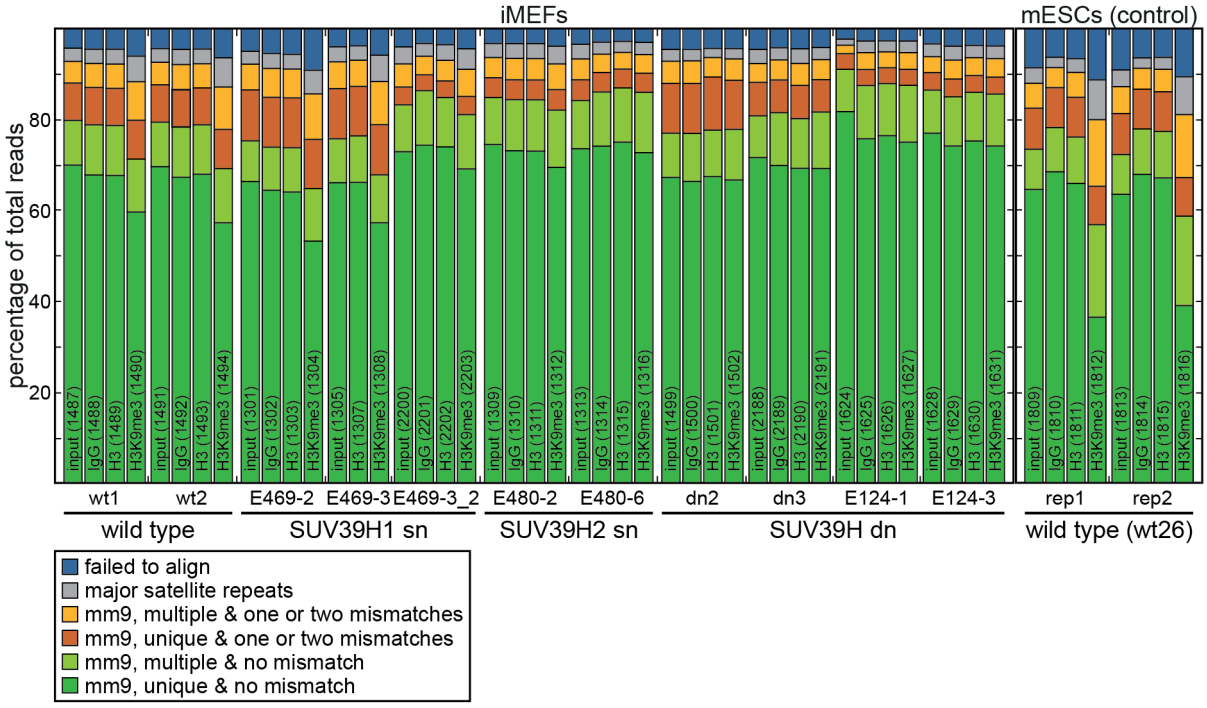

Figure 46: Mapping results of all samples left after QC. Total read numbers and coverage are listed in **Supplementary Table S2**. Notably, H3K9me3 ChIP samples from wild type (iMEFs and mESCs) and most SUV39H1 sn cell lines have an increased fraction of major satellite mapping reads (grey), multiple mapping reads (yellow and light green) as well as unaligned reads (blue) compared to controls.

ChIP-seq data analyses. The data was analyzed according to the overview below (**Fig. 47**). The coverage was calculated using a Perl script and all other calculations were done using `R`. Briefly, reads that overlap with feature i were counted using samtools (`multiBamCov`) and normalized to the average sample coverage c and the feature length k_i.

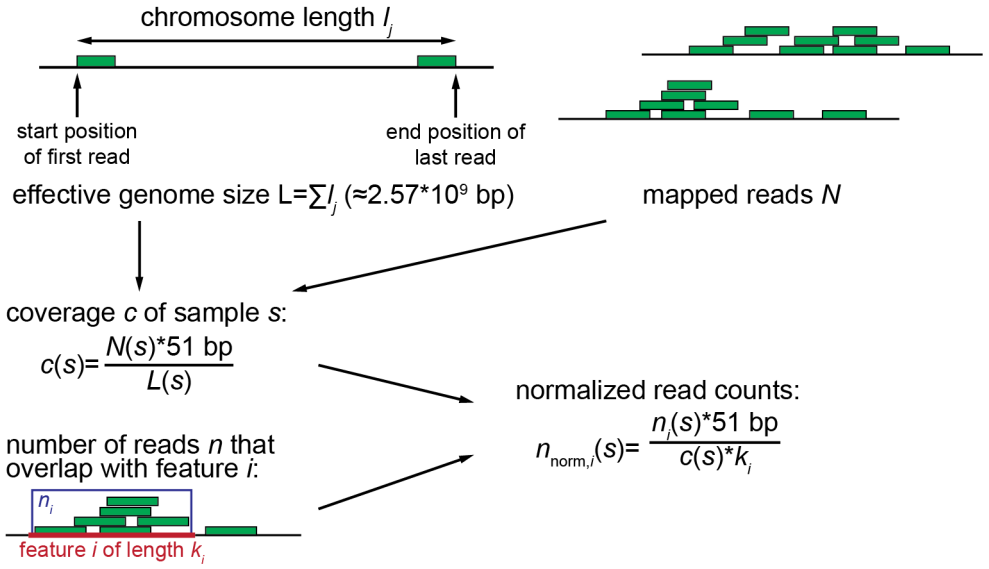

Figure 47: Calculation of sample coverage and normalized read counts per feature. The coverage of each sample was determined from the effective chromosome length and the total number of mapped reads (multiple mapping allowing up to two mismatches). Reads overlapping with features were counted using `multiBamCov` (samtools). Feature lengths were taken from annotation bed files.

The resulting normalized read counts $n_{\mathrm{norm},i}$ and the background contribution b determined with MCORE ((Molitor *et al.* 2017)) were then used to calculate the enrichment of a feature. To be able to compare different samples, the enrichment was normalized with $(1-b)$ yielding a genomic average of enrichment $E = 1$ for all samples:

$$E_i = \frac{n_{\mathrm{norm},i}(\mathrm{IP}) - b \cdot n_{\mathrm{norm},i}(\mathrm{IgG})}{n_{\mathrm{norm},i}(\mathrm{input})} \cdot \frac{1}{1-b} \tag{10}$$

Features with $n_{\mathrm{norm},i} < 0.3$ (i. e. feature input covered less than 30% of the average genome-wide input coverage) were considered "unmappable" and excluded. For all analyses, the average of cell lines with the same genotype was analyzed (see **Supplementary Table S2** for a summary). Features with $E_i > 3$ (i. e. threefold enrichment over genomic background) were considered to be highly methylated. To analyze the methylated subset, features with $E_i(\mathrm{iMEFwt}) > 3$ were selected and the fold enrichment $E_{i,\mathrm{fold}}(\mathrm{sample}) = E_i(\mathrm{sample})/E_i(\mathrm{iMEFwt})$ was calculated for the SUV39H sn, dn and mESC wt samples. For the RefSeq genes, genes with identical genomic start and end positions (i. e. splice variants) were only kept once as they are the same on the DNA

level.

IV.2.12 RNA sequencing

All RNA-seq results described in this study were deduced from sequencing strand-specific libraries of long (>200 nt) total RNA libraries. An overview of the workflow is given in the Appendix (**Supplementary Fig. S2**).

RNA isolation. $3 \times 10^5 - 1.5 \times 10^6$ cells were dissolved in a 1:1 mixture of iT buffer (**Table 10**) and glycerol and stored at -80°C until all samples were collected and processed together. After thawing the cells, they were washed twice with ice-cold PBS before RNA isolation using the miRNeasy Mini kit (#217004, Qiagen, Netherlands). RNA isolation was carried out according to the manufacturer's protocol and both the miRNA-enriched fraction and the total RNA fraction (>200 nt) were purified separately. Subsequently, residual genomic DNA was removed by DNaseI. Therefore, the RNA (in RNase-free water) was mixed with RQ1 DNase reaction buffer (1×), 1 U/µl ribolock (ThermoFisher Scientific, USA) and 1 U RQ1 DNase (Promega, USA) per µg of RNA and incubated at 37°C for 30 min. Subsequently, the RNA was precipitated by adding sodium acetate to a final concentration of 100 mM, 0.4 mg/ml glycogen and three volumes of 100% ethanol. After incubation at -20°C for at least 15 min, RNA was pelleted by centrifugation in a benchtop centrifuge at full speed and 4°C. After washing the pellet once with 70% ethanol, the RNA was dissolved in RNase-free water, the concentration was determined by Qubit (RNA high sensitivity, ThermoFisher Scientific, USA) and the RNA integrity (long RNAs) was checked on a 2% w/v agarose E-gel (ThermoFisher Scientific, USA). Therefor, ∼200 ng RNA were diluted in 10 µl water and mixed 1:1 with formamid. After heating to 65°C for 5 min, the samples were immediately stored on ice, then loaded onto the gel and separated for 10 min (program 7).

Library preparation. Libraries were only prepared for the long RNA fraction. First, ribosomal RNA was removed using the Ribo-Zero Magnetic Gold Kit (Epicentre Biotechnologies, USA) with 2 µg RNA input according to the manufacturer's protocol. After measuring the concentration with Qubit (RNA high sensitivity, ThermoFisher Scientific, USA), 20-40 ng RNA were used for the strand-specific library preparation with the NEBNext RNA library prep kit for Illumina (#E7420, New England Biolabs, USA) following the provided protocol. The PCR library enrichment was done with 13 cycles. Last, library size and quality were checked on a bioanalyzer chip (DNA1000 and High sensitivity, Agilent Technologies, USA). The libraries were sequenced and demultiplexed by the DKFZ core facility on a HiSeq 2000 machine (Illumina, USA).

Initial data processing. All samples were mapped to the mouse major satellite consensus sequence (Lehnertz *et al.* 2003) without mismatches using Bowtie (Langmead *et al.* 2009). All unmapped reads were then mapped to the mm9 assembly of the reference genome with random assignment of multiple mapping reads and allowing up to two

mismatches using the spliced read aligner STAR (Dobin *et al.* 2013). A STAR index file for the mm9 assembly of the reference genome was generated beforehand with an mm9 fasta file and a gene annotation file (gtf) that was also used for the mapping (providing annotated intron/exon boundaries). The gene annotation file contained the mm9 annotation of the RefSeq genes (whole genes) and was downloaded from the UCSC table browser (Karolchik *et al.* 2004). The sorted output bam files were indexed using samtools (Li & Durbin 2009).

RNA-seq data analyses. Reads overlapping with certain features (i. e. GSATs, HET6387 regions or genes) were counted using `multiBamCov` (samtools) with the options considering split reads and using the strand information. Features that were excluded from the ChIP-seq analyses due to limited mappability (see above) were also excluded from the RNA-seq analyses for consistency. Despite its limitations (Wagner *et al.* 2012), the RPKM measure of read density (Mortazavi *et al.* 2008) was calculated for the features listed above and used as an estimate of transcript abundance. For a more accurate quantification of log2-fold changes of RNA-seq samples from SUV39H sn or dn cell lines compared to wt iMEFs, the `R` tool DESeq2 (Love *et al.* 2014) was used. Importantly, the DESeq2 analysis was done for the full set of (mappable) RefSeq genes ($n = 26,036$) to ensure accurate normalization. The partitioning in different subsets of genes was done in the downstream analysis.

References

L. Aagaard, G. Laible, P. Selenko, M. Schmid, R. Dorn, G. Schotta, S. Kuhfittig, A. Wolf, A. Lebersorger, P. B. Singh, G. Reuter & T. Jenuwein. 1999. Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. *EMBO J.*, 18(7):1923–1938.

S. Ait-Si-Ali, V. Guasconi, L. Fritsch, H. Yahi, R. Sekhri, I. Naguibneva, P. Robin, F. Cabon, A. Polesskaya & A. Harel-Bellan. 2004. A Suv39h-dependent mechanism for silencing S-phase genes in differentiating but not in cycling cells. *EMBO J.*, 23(3):605–615.

C. Alabert, J. C. Bukowski-Wills, S. B. Lee, G. Kustatscher, K. Nakamura, F. de Lima Alves, P. Menard, J. Mejlvang, J. Rappsilber & A. Groth. 2014. Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nat. Cell Biol.*, 16(3):281–293.

B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts & J. D. Watson. 1994. *Molecular biology of the cell.* Garland Publishing, New York and London, 3rd edition.

J. Arand, D. Spieler, T. Karius, M. R. Branco, D. Meilinger, A. Meissner, T. Jenuwein, G. Xu, H. Leonhardt, V. Wolf & J. Walter. 2012. In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS Genet.*, 8(6):e1002750.

K. Ayyanathan, M. S. Lechner, P. Bell, G. G. Maul, D. C. Schultz, Y. Yamada, K. Tanaka, K. Torigoe & F. J. Rauscher. 2003. Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev.*, 17(15):1855–1869.

M. Baker. 2011. Making sense of chromatin states. *Nat. Methods*, 8(9):717–722.

A. J. Bannister & T. Kouzarides. 2011. Regulation of chromatin by histone modifications. *Cell Res.*, 21(3):381–395.

A. J. Bannister, P. Zegerman, J. F. Partridge, E. A. Miska, J. O. Thomas, R. C. Allshire & T. Kouzarides. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, 410(6824):120–124.

A. Barski, S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev & K. Zhao. 2007. High-resolution profiling of histone methylations in the human genome. *Cell*, 129(4):823–837.

B. Bartkowiak, P. Liu, H. P. Phatnani, N. J. Fuda, J. J. Cooper, D. H. Price, K. Adelman, J. T. Lis & A. L. Greenleaf. 2010. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev.*, 24(20):2303–2316.

A. Basavapathruni, J. Gureasko, M. Porter Scott, W. Hermans, A. Godbole, P. A. Leland, P. A. Boriack-Sjodin, T. J. Wigle, R. A. Copeland & T. V. Riera. 2016. Characterization of the Enzymatic Activity of SETDB1 and Its 1:1 Complex with ATF7IP. *Biochemistry*, 55(11): 1645–1651.

M. Baumann, J. Pontiller & W. Ernst. 2010. Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. *Mol. Biotechnol.*, 45 (3):241–247.

J. S. Becker, D. Nicetto & K. S. Zaret. 2016. H3K9-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet.*, 32(1):29–41.

R. R. Beerli, D. J. Segal, B. Dreier & C. F. Barbas. 1998. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U.S.A.*, 95(25):14628–14633.

A. Benard, I. J. Goossens-Beumer, A. Q. van Hoesel, W. de Graaf, H. Horati, H. Putter, E. C. Zeestraten, C. J. van de Velde & P. J. Kuppen. 2014. Histone trimethylation at H3K4, H3K9 and H4K20 correlates with patient survival and tumor recurrence in early-stage colon cancer. *BMC Cancer*, 14:531.

C. Berens & W. Hillen. 2003. Gene regulation by tetracyclines. Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes. *Eur. J. Biochem.*, 270(15): 3109–3121.

S. Berry, C. Dean & M. Howard. 2017. Slow Chromatin Dynamics Allow Polycomb Target Genes to Filter Fluctuations in Transcription Factor Activity. *Cell Syst*, 4(4):445–457.

L. Bintu, J. Yong, Y. E. Antebi, K. McCue, Y. Kazuki, N. Uno, M. Oshimura & M. B. Elowitz. 2016. Dynamics of epigenetic regulation at the single-cell level. *Science*, 351(6274):720–724.

A. Bird. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev.*, 16(1):6–21.

J. Blau, H. Xiao, S. McCracken, P. O'Hare, J. Greenblatt & D. Bentley. 1996. Three functional classes of transcriptional activation domain. *Mol. Cell. Biol.*, 16(5):2044–2055.

J. Boch & U. Bonas. 2010. Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol*, 48:419–436.

J. Boch, H. Scholze, S. Schornack, A. Landgraf, S. Hahn, S. Kay, T. Lahaye, A. Nickstadt & U. Bonas. 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, 326(5959):1509–1512.

A. N. Boettiger, B. Bintu, J. R. Moffitt, S. Wang, B. J. Beliveau, G. Fudenberg, M. Imakaev, L. A. Mirny, C. T. Wu & X. Zhuang. 2016. Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature*, 529(7586):418–422.

W. E. Briley, M. H. Bondy, P. S. Randeria, T. J. Dupper & C. A. Mirkin. 2015. Quantification and real-time tracking of RNA in live cells using Sticky-flares. *Proc. Natl. Acad. Sci. U.S.A.*, 112(31):9591–9595.

L. Brueckner, J. van Arensbergen, W. Akhtar, L. Pagie & B. van Steensel. 2016. High-throughput assessment of context-dependent effects of chromatin proteins. *Epigenetics Chromatin*, 9:43.

L. J. Bugaj, A. T. Choksi, C. K. Mesuda, R. S. Kane & D. V. Schaffer. 2013. Optogenetic protein clustering and signaling activation in mammalian cells. *Nat. Methods*, 10(3):249–252.

A. Bulut-Karslioglu, V. Perrera, M. Scaranaro, I. A. de la Rosa-Velazquez, S. van de Nobelen, N. Shukeir, J. Popow, B. Gerle, S. Opravil, M. Pagani, S. Meidhof, T. Brabletz, T. Manke, M. Lachner & T. Jenuwein. 2012. A transcription factor-based mechanism for mouse heterochromatin formation. *Nat. Struct. Mol. Biol.*, 19(10):1023–1030.

A. Bulut-Karslioglu, I. A. De La Rosa-Velazquez, F. Ramirez, M. Barenboim, M. Onishi-Seebacher, J. Arand, C. Galan, G. E. Winter, B. Engist, B. Gerle, R. J. O'Sullivan, J. H. Martens, J. Walter, T. Manke, M. Lachner & T. Jenuwein. 2014. Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol. Cell*, 55(2):277–290.

S. Buratowski. 2009. Progression through the RNA polymerase II CTD cycle. *Mol. Cell*, 36(4): 541–546.

A. E. Carpenter, S. Memedula, M. J. Plutz & A. S. Belmont. 2005. Common effects of acidic activators on large-scale chromatin structure and transcription. *Mol. Cell. Biol.*, 25(3):958–968.

M. J. Carrozza, B. Li, L. Florens, T. Suganuma, S. K. Swanson, K. K. Lee, W. J. Shia, S. Anderson, J. Yates, M. P. Washburn & J. L. Workman. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*, 123(4):581–592.

M. Caudron-Herger, T. Pankert, J. Seiler, A. Nemeth, R. Voit, I. Grummt & K. Rippe. 2015. Alu element-containing RNAs maintain nucleolar structure and function. *EMBO J.*, 34(22): 2758–2774.

G. Chalepakis, J. Wijnholds & P. Gruss. 1994. Pax-3-DNA interaction: flexibility in the DNA binding and induction of DNA conformational changes by paired domains. *Nucleic Acids Res.*, 22(15):3131–3137.

B. Chen, L. A. Gilbert, B. A. Cimini, J. Schnitzbauer, W. Zhang, G. W. Li, J. Park, E. H. Blackburn, J. S. Weissman, L. S. Qi & B. Huang. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, 155(7):1479–1491.

B. Cheng, T. Li, P. B. Rahl, T. E. Adamson, N. B. Loudas, J. Guo, K. Varzavand, J. J. Cooper, X. Hu, A. Gnatt, R. A. Young & D. H. Price. 2012. Functional association of Gdown1 with RNA polymerase II poised on human genes. *Mol. Cell*, 45(1):38–50.

S. R. Choudhury, Y. Cui, A. Narayanan, D. P. Gilley, N. Huda, C. L. Lo, F. C. Zhou, D. Yernool & J. Irudayaraj. 2016. Optogenetic regulation of site-specific subtelomeric DNA methylation. *Oncotarget*, 7(31):50380–50391.

I. Chung, H. Leonhardt & K. Rippe. 2011. De novo assembly of a PML nuclear subcompartment occurs through multiple pathways and induces telomere elongation. *J. Cell. Sci.*, 124(Pt 21): 3603–3618.

R. C. Conaway & J. W. Conaway. 2011. Origins and activity of the Mediator complex. *Semin. Cell Dev. Biol.*, 22(7):729–734.

L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini & F. Zhang. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339(6121):819–823.

J. L. Corden. 1990. Tails of RNA polymerase II. *Trends Biochem. Sci.*, 15(10):383–387.

M. Corrales, A. Rosado, R. Cortini, J. van Arensbergen, B. van Steensel & G. J. Filion. 2017. Clustering of Drosophila housekeeping promoters facilitates their expression. *Genome Res.*, 27(7):1153–1161.

G. N. Corry, M. J. Hendzel & D. A. Underhill. 2008. Subnuclear localization and mobility are key indicators of PAX3 dysfunction in Waardenburg syndrome. *Hum. Mol. Genet.*, 17(12): 1825–1837.

L. T. M. Dao, A. O. Galindo-Albarran, J. A. Castro-Mondragon, C. Andrieu-Soler, A. Medina-Rivera, C. Souaid, G. Charbonnier, A. Griffon, L. Vanhille, T. Stephen, J. Alomairi, D. Martin, M. Torres, N. Fernandez, E. Soler, J. van Helden, D. Puthier & S. Spicuglia. 2017. Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet.*, 49(7):1073–1081.

X. Darzacq, J. Yao, D. R. Larson, S. Z. Causse, L. Bosanac, V. de Turris, V. M. Ruda, T. Lionnet, D. Zenklusen, B. Guglielmi, R. Tjian & R. H. Singer. 2009. Imaging transcription in living cells. *Annu Rev Biophys*, 38:173–196.

A. Dhayalan, R. Tamas, I. Bock, A. Tattermusch, E. Dimitrova, S. Kudithipudi, S. Ragozin & A. Jeltsch. 2011. The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Hum. Mol. Genet.*, 20(11):2195–2203.

Y. Diao, R. Fang, B. Li, Z. Meng, J. Yu, Y. Qiu, K. C. Lin, H. Huang, T. Liu, R. J. Marina, I. Jung, Y. Shen, K. L. Guan & B. Ren. 2017. A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. *Nat. Methods*, 14(6):629–635.

A. Diaz, A. Nellore & J. S. Song. 2012. CHANCE: comprehensive software for quality control and validation of ChIP-seq data. *Genome Biol.*, 13(10):R98.

P. Dimitri, N. Corradini, F. Rossi & F. Verni. 2005. The paradox of functional heterochromatin. *Bioessays*, 27(1):29–41.

A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson & T. R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1): 15–21.

I. B. Dodd, M. A. Micheelsen, K. Sneppen & G. Thon. 2007. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell*, 129(4):813–822.

J. E. Dodge, Y. K. Kang, H. Beppu, H. Lei & E. Li. 2004. Histone H3-K9 methyltransferase ESET is essential for early development. *Mol. Cell. Biol.*, 24(6):2478–2486.

K. B. Dong, I. A. Maksakova, F. Mohn, D. Leung, R. Appanah, S. Lee, H. W. Yang, L. L. Lam, D. L. Mager, D. Schubeler, M. Tachibana, Y. Shinkai & M. C. Lorincz. 2008. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J.*, 27(20):2691–2701.

I. du Chene, E. Basyuk, Y. L. Lin, R. Triboulet, A. Knezevich, C. Chable-Bessia, C. Mettling, V. Baillat, J. Reynes, P. Corbeau, E. Bertrand, A. Marcello, S. Emiliani, R. Kiernan & M. Benkirane. 2007. Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *EMBO J.*, 26(2):424–435.

J. W. Edmunds, L. C. Mahadevan & A. L. Clayton. 2008. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J.*, 27(2): 406–420.

S. Egloff, M. Dienstbier & S. Murphy. Jul 2012. Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends Genet.*, 28(7):333–341.

S. C. Elgin & G. Reuter. 2013. Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harb Perspect Biol*, 5(8):a017780.

G. C. Ellis-Davies. 2007. Caged compounds: photorelease technology for control of cellular chemistry and physiology. *Nat. Methods*, 4(8):619–628.

F. Erdel & E. C. Greene. 2016. Generalized nucleation and looping model for epigenetic memory of histone modifications. *Proc. Natl. Acad. Sci. U.S.A.*, 113(29):E4180–4189.

T. Y. Erkina & A. M. Erkine. 2016. Nucleosome distortion as a possible mechanism of transcription activation domain function. *Epigenetics Chromatin*, 9:40.

J. Ernst & M. Kellis. 2010. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.*, 28(8):817–825.

C. Fabrega, V. Shen, S. Shuman & C. D. Lima. 2003. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Mol. Cell*, 11 (6):1549–1561.

A. Fieck, D. L. Wyborski & J. M. Short. 1992. Modifications of the E.coli Lac repressor for expression in eukaryotic cells: effects of nuclear signal sequences on protein activity and nuclear accumulation. *Nucleic Acids Res.*, 20(7):1785–1791.

B. Fierz, C. Chatterjee, R. K. McGinty, M. Bar-Dagan, D. P. Raleigh & T. W. Muir. 2011. Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nat. Chem. Biol.*, 7(2):113–119.

G. S. Filonov, K. D. Piatkevich, L. M. Ting, J. Zhang, K. Kim & V. V. Verkhusha. 2011. Bright and stable near-infrared fluorescent protein for in vivo imaging. *Nat. Biotechnol.*, 29 (8):757–761.

R. Firestein, X. Cui, P. Huie & M. L. Cleary. 2000. Set domain-dependent regulation of transcriptional silencing and growth control by SUV39H1, a mammalian ortholog of Drosophila Su(var)3-9. *Mol. Cell. Biol.*, 20(13):4900–4909.

L. Fritsch, P. Robin, J. R. Mathieu, M. Souidi, H. Hinaux, C. Rougeulle, A. Harel-Bellan, M. Ameyar-Zazoua & S. Ait-Si-Ali. 2010. A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. *Mol. Cell*, 37 (1):46–56.

G. Fuchs, Y. Voichek, S. Benjamin, S. Gilad, I. Amit & M. Oren. 2014. 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biol.*, 15(5):R69.

F. Fuks. 2005. DNA methylation and histone modifications: teaming up to silence genes. *Curr. Opin. Genet. Dev.*, 15(5):490–495.

F. Fuks, P. J. Hurd, R. Deplus & T. Kouzarides. 2003a. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.*, 31(9):2305–2312.

F. Fuks, P. J. Hurd, D. Wolf, X. Nan, A. P. Bird & T. Kouzarides. 2003b. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J. Biol. Chem.*, 278 (6):4035–4040.

X. Gao, J. C. Tsang, F. Gaba, D. Wu, L. Lu & P. Liu. 2014. Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. *Nucleic Acids Res.*, 42(20):e155.

P. T. Georgel, R. A. Horowitz-Scherer, N. Adkins, C. L. Woodcock, P. A. Wade & J. C. Hansen. 2003. Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J. Biol. Chem.*, 278(34):32181–32188.

A. Ghamari, M. P. van de Corput, S. Thongjuea, W. A. van Cappellen, W. van Ijcken, J. van Haren, E. Soler, D. Eick, B. Lenhard & F. G. Grosveld. 2013. In vivo live imaging of RNA polymerase II transcription factories in primary cells. *Genes Dev.*, 27(7):767–777.

R. P. Ghosh, R. A. Horowitz-Scherer, T. Nikitina, L. S. Shlyakhtenko & C. L. Woodcock. 2010. MeCP2 binds cooperatively to its substrate and competes with histone H1 for chromatin binding sites. *Mol. Cell. Biol.*, 30(19):4656–4670.

D. A. Gilchrist, G. Dos Santos, D. C. Fargo, B. Xie, Y. Gao, L. Li & K. Adelman. 2010. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell*, 143(4):540–551.

J. L. Goodier, P. K. Mandal, L. Zhang & H. H. Kazazian. 2010. Discrete subcellular partitioning of human retrotransposon RNAs despite a common mechanism of genome insertion. *Hum. Mol. Genet.*, 19(9):1712–1725.

S. M. Gorisch, M. Wachsmuth, K. F. Toth, P. Lichter & K. Rippe. 2005. Histone acetylation increases chromatin accessibility. *J. Cell. Sci.*, 118(Pt 24):5825–5834.

M. Gossen & H. Bujard. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.*, 89(12):5547–5551.

V. Grishkevich & I. Yanai. 2014. Gene length and expression level shape genomic novelties. *Genome Res.*, 24(9):1497–1503.

M. Guenatri, D. Bailly, C. Maison & G. Almouzni. 2004. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.*, 166(4):493–505.

M. G. Guenther, S. S. Levine, L. A. Boyer, R. Jaenisch & R. A. Young. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, 130(1):77–88.

N. A. Hathaway, O. Bell, C. Hodges, E. L. Miller, D. S. Neel & G. R. Crabtree. 2012. Dynamics and memory of heterochromatin in living cells. *Cell*, 149(7):1447–1460.

M. Heidemann, C. Hintermair, K. Voss & D. Eick. 2013. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim. Biophys. Acta*, 1829(1):55–62.

S. Heinz, C. E. Romanoski, C. Benner & C. K. Glass. Mar 2015. The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.*, 16(3):144–154.

E. Heitz. 1928. *Das Heterochromatin der Moose*, volume 69.

T. Henriques, D. A. Gilchrist, S. Nechaev, M. Bern, G. W. Muse, A. Burkholder, D. C. Fargo & K. Adelman. 2013. Stable pausing by RNA polymerase II provides an opportunity to target and integrate regulatory signals. *Mol. Cell*, 52(4):517–528.

H. Herrmann & U. Aebi. 2004. Intermediate filaments: molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. *Annual review of biochemistry*, 73:749–89.

H. Hirai, T. Tani & N. Kikyo. 2010. Structure and functions of powerful transactivators: VP16, MyoD and FoxA. *Int. J. Dev. Biol.*, 54(11-12):1589–1596.

F. C. Holstege, U. Fiedler & H. T. Timmers. 1997. Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO J.*, 16(24):7468–7480.

L. Hong, G. P. Schroth, H. R. Matthews, P. Yau & E. M. Bradbury. 1993. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J. Biol. Chem.*, 268(1):305–314.

P. Horvath & R. Barrangou. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science*, 327(5962):167–170.

W. Horz & W. Altenburger. 1981. Nucleotide sequence of mouse satellite DNA. *Nucleic Acids Res.*, 9(3):683–696.

P. D. Hsu, E. S. Lander & F. Zhang. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 157(6):1262–1278.

K. L. Huisinga, B. Brower-Toland & S. C. Elgin. 2006. The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma*, 115(2):110–122.

J. P. Jackson, A. M. Lindroth, X. Cao & S. E. Jacobsen. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature*, 416(6880):556–560.

R. H. Jacobson, A. G. Ladurner, D. S. King & R. Tjian. 2000. Structure and function of a human TAFII250 double bromodomain module. *Science*, 288(5470):1422–1425.

S. M. Janicki, T. Tsukamoto, S. E. Salghetti, W. P. Tansey, R. Sachidanandam, K. V. Prasanth, T. Ried, Y. Shav-Tal, E. Bertrand, R. H. Singer & D. L. Spector. 2004. From silencing to gene expression: real-time analysis in single cells. *Cell*, 116(5):683–698.

T. Jegou, I. Chung, G. Heuvelman, M. Wachsmuth, S. M. Gorisch, K. M. Greulich-Bode, P. Boukamp, P. Lichter & K. Rippe. 2009. Dynamics of telomeres and promyelocytic leukemia nuclear bodies in a telomerase-negative human cell line. *Mol. Biol. Cell*, 20(7):2070–2082.

W. L. Johnson, W. T. Yewdell, J. C. Bell, S. M. McNulty, Z. Duda, R. J. O'Neill, B. A. Sullivan & A. F. Straight. 2017. RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife*, 6.

I. Jonkers & J. T. Lis. 2015. Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.*, 16(3):167–177.

D. Karolchik, A. S. Hinrichs, T. S. Furey, K. M. Roskin, C. W. Sugnet, D. Haussler & W. J. Kent. 2004. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.*, 32(Database issue):D493–496.

F. Kawano, H. Suzuki, A. Furuya & M. Sato. 2015. Engineered pairs of distinct photoswitches for optogenetic control of cellular proteins. *Nat Commun*, 6:6256.

M. J. Kennedy, R. M. Hughes, L. A. Peteya, J. W. Schwartz, M. D. Ehlers & C. L. Tucker. 2010. Rapid blue-light-mediated induction of protein interactions in living cells. *Nat. Methods*, 7 (12):973–975.

J. Kim, J. Daniel, A. Espejo, A. Lake, M. Krishna, L. Xia, Y. Zhang & M. T. Bedford. 2006. Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep.*, 7(4): 397–403.

K. O. Kizer, H. P. Phatnani, Y. Shibata, H. Hall, A. L. Greenleaf & B. D. Strahl. 2005. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol. Cell. Biol.*, 25(8):3305–3316.

S. Konermann, M. D. Brigham, A. Trevino, P. D. Hsu, M. Heidenreich, L. Cong, R. J. Platt, D. A. Scott, G. M. Church & F. Zhang. 2013. Optical control of mammalian endogenous transcription and epigenetic states. *Nature*, 500(7463):472–476.

S. Konermann, M. D. Brigham, A. E. Trevino, J. Joung, O. O. Abudayyeh, C. Barcena, P. D. Hsu, N. Habib, J. S. Gootenberg, H. Nishimasu, O. Nureki & F. Zhang. 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, 517(7536): 583–588.

T. Kouzarides. 2002. Histone methylation in transcriptional control. *Curr. Opin. Genet. Dev.*, 12(2):198–209.

T. Kouzarides. 2007. Chromatin modifications and their function. *Cell*, 128(4):693–705.

I. M. Krouwels, K. Wiesmeijer, T. E. Abraham, C. Molenaar, N. P. Verwoerd, H. J. Tanke & R. W. Dirks. 2005. A glue for heterochromatin maintenance: stable SUV39H1 binding to heterochromatin is reinforced by the SET domain. *J. Cell Biol.*, 170(4):537–549.

S. Kudithipudi, M. K. Schuhmacher, A. F. Kebede & A. Jeltsch. 2017. The SUV39H1 Protein Lysine Methyltransferase Methylates Chromatin Proteins Involved in Heterochromatin Formation and VDJ Recombination. *ACS Chem. Biol.*, 12(4):958–968.

G. Kungulovski, I. Kycia, R. Tamas, R. Z. Jurkowska, S. Kudithipudi, C. Henry, R. Reinhardt, P. Labhart & A. Jeltsch. 2014. Application of histone modification-specific interaction domains as an alternative to antibodies. *Genome Res.*, 24(11):1842–1853.

G. Kungulovski, S. Nunna, M. Thomas, U. M. Zanger, R. Reinhardt & A. Jeltsch. 2015. Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. *Epigenetics Chromatin*, 8:12.

H. Kwak, N. J. Fuda, L. J. Core & J. T. Lis. 2013. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science*, 339(6122):950–953.

B. Langmead, C. Trapnell, M. Pop & S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, 10(3):R25.

D. R. Larson, C. Fritzsch, L. Sun, X. Meng, D. S. Lawrence & R. H. Singer. 2013. Direct observation of frequency modulated transcription in single cells using light activation. *Elife*, 2:e00750.

M. Lawrence, S. Daujat & R. Schneider. 2016. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet.*, 32(1):42–56.

T. Lederer, M. Takahashi & W. Hillen. 1995. Thermodynamic analysis of tetracycline-mediated induction of Tet repressor by a quantitative methylation protection assay. *Anal. Biochem.*, 232(2):190–196.

B. Lehnertz, Y. Ueda, A. A. Derijck, U. Braunschweig, L. Perez-Burgos, S. Kubicek, T. Chen, E. Li, T. Jenuwein & A. H. Peters. 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.*, 13(14):1192–1200.

C. Lerin, J. T. Rodgers, D. E. Kalume, S. H. Kim, A. Pandey & P. Puigserver. 2006. GCN5 acetyltransferase complex controls glucose metabolism through transcriptional repression of PGC-1alpha. *Cell Metab.*, 3(6):429–438.

G. LeRoy, B. Rickards & S. J. Flint. 2008. The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol. Cell*, 30(1):51–60.

J. D. Lewis, R. R. Meehan, W. J. Henzel, I. Maurer-Fogy, P. Jeppesen, F. Klein & A. Bird. 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, 69(6):905–914.

G. Li, X. Ruan, R. K. Auerbach, K. S. Sandhu, M. Zheng, P. Wang, H. M. Poh, Y. Goh, J. Lim, J. Zhang, H. S. Sim, S. Q. Peh, F. H. Mulawadi, C. T. Ong, Y. L. Orlov, S. Hong, Z. Zhang, S. Landt, D. Raha, G. Euskirchen, C. L. Wei, W. Ge, H. Wang, C. Davis, K. I. Fisher-Aylor, A. Mortazavi, M. Gerstein, T. Gingeras, B. Wold, Y. Sun, M. J. Fullwood, E. Cheung, E. Liu, W. K. Sung, M. Snyder & Y. Ruan. 2012. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell*, 148(1-2):84–98.

H. Li & R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14):1754–1760.

H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis & R. Durbin. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25 (16):2078–2079.

J. Li, Q. Lin, H. G. Yoon, Z. Q. Huang, B. D. Strahl, C. D. Allis & J. Wong. 2002. Involvement of histone methylation and phosphorylation in regulation of transcription by thyroid hormone receptor. *Mol. Cell. Biol.*, 22(16):5688–5697.

J. Li, Z. Li, J. Ruan, C. Xu, Y. Tong, P. W. Pan, W. Tempel, L. Crombet, J. Min & J. Zang. 2011. Structural basis for specific binding of human MPP8 chromodomain to histone H3 methylated at lysine 9. *PLoS ONE*, 6(10):e25104.

G. Liang, J. C. Lin, V. Wei, C. Yoo, J. C. Cheng, C. T. Nguyen, D. J. Weisenberger, G. Egger, D. Takai, F. A. Gonzales & P. A. Jones. 2004. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc. Natl. Acad. Sci. U.S.A.*, 101(19):7357–7362.

H. Liu, X. Yu, K. Li, J. Klejnot, H. Yang, D. Lisiero & C. Lin. 2008. Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in Arabidopsis. *Science*, 322 (5907):1535–1539.

M. I. Love, W. Huber & S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, 15(12):550.

C. Lukas, J. Falck, J. Bartkova, J. Bartek & J. Lukas. 2003. Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage. *Nat. Cell Biol.*, 5(3):255–260.

C. Maison, D. Bailly, D. Roche, R. Montes de Oca, A. V. Probst, I. Vassias, F. Dingli, B. Lombard, D. Loew, J. P. Quivy & G. Almouzni. 2011. SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin. *Nat. Genet.*, 43(3):220–227.

C. Maison, D. Bailly, J. P. Quivy & G. Almouzni. 2016. The methyltransferase Suv39h1 links the SUMO pathway to HP1alpha marking at pericentric heterochromatin. *Nat Commun*, 7: 12224.

P. Mali, L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville & G. M. Church. 2013. RNA-guided human genome engineering via Cas9. *Science*, 339(6121):823–826.

L. Manuelidis. 1981. Consensus sequence of mouse satellite DNA indicates it is derived from tandem 116 basepair repeats. *FEBS Lett.*, 129(1):25–28.

J. H. Martens, R. J. O'Sullivan, U. Braunschweig, S. Opravil, M. Radolf, P. Steinlein & T. Jenuwein. 2005. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.*, 24(4):800–812.

A. Meinhart, T. Kamenski, S. Hoeppner, S. Baumli & P. Cramer. 2005. A structural perspective of CTD function. *Genes Dev.*, 19(12):1401–1415.

M. Melcher, M. Schmid, L. Aagaard, P. Selenko, G. Laible & T. Jenuwein. 2000. Structure-function analysis of SUV39H1 reveals a dominant role in heterochromatin organization, chromosome segregation, and mitotic progression. *Mol. Cell. Biol.*, 20(10):3728–3741.

J. Molitor, J. P. Mallm, K. Rippe & F. Erdel. 2017. Retrieving Chromatin Patterns from Deep Sequencing Data Using Correlation Functions. *Biophys. J.*, 112(3):473–490.

A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer & B. Wold. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods*, 5(7):621–628.

L. B. Motta-Mena, A. Reade, M. J. Mallory, S. Glantz, O. D. Weiner, K. W. Lynch & K. H. Gardner. 2014. An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nat. Chem. Biol.*, 10(3):196–202.

F. Mueller, P. Wach & J. G. McNally. 2008. Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching. *Biophys. J.*, 94(8):3323–3339.

F. Mueller, D. Mazza, T. J. Stasevich & J. G. McNally. 2010. FRAP and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know? *Curr. Opin. Cell Biol.*, 22 (3):403–411.

H. J. Muller. 1930. Types of visible variations induced by x-rays in drosophila. *Journal of Genetics*, 22(3):299–334.

K. P. Muller, F. Erdel, M. Caudron-Herger, C. Marth, B. D. Fodor, M. Richter, M. Scaranaro, J. Beaudouin, M. Wachsmuth & K. Rippe. 2009. Multiscale analysis of dynamics and interactions of heterochromatin protein 1 by fluorescence fluctuation microscopy. *Biophys. J.*, 97 (11):2876–2885.

K. Muller-Ott, F. Erdel, A. Matveeva, J. P. Mallm, A. Rademacher, M. Hahn, C. Bauer, Q. Zhang, S. Kaltofen, G. Schotta, T. Hofer & K. Rippe. 2014. Specificity, propagation, and memory of pericentric heterochromatin. *Mol. Syst. Biol.*, 10:746.

C. A. Musselman, M. E. Lalonde, J. Cote & T. G. Kutateladze. 2012. Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.*, 19(12):1218–1227.

J. Nakayama, J. C. Rice, B. D. Strahl, C. D. Allis & S. I. Grewal. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science*, 292(5514): 110–113.

H. H. Ng, F. Robert, R. A. Young & K. Struhl. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell*, 11(3):709–719.

M. Ni, J. M. Tepperman & P. H. Quail. 1999. Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature*, 400(6746):781–784.

S. J. Nielsen, R. Schneider, U. M. Bauer, A. J. Bannister, A. Morrison, D. O'Carroll, R. Firestein, M. Cleary, T. Jenuwein, R. E. Herrera & T. Kouzarides. 2001. Rb targets histone H3 methylation and HP1 to promoters. *Nature*, 412(6846):561–565.

D. Niopek, D. Benzinger, J. Roensch, T. Draebing, P. Wehler, R. Eils & B. Di Ventura. Jul 2014. Engineering light-inducible nuclear localization signals for precise spatiotemporal control of protein dynamics in living cells. *Nat Commun*, 5:4404.

D. Niopek, P. Wehler, J. Roensch, R. Eils & B. Di Ventura. 2016. Optogenetic control of nuclear protein export. *Nat Commun*, 7:10624.

N. Nocetti & I. Whitehouse. 2016. Nucleosome repositioning underlies dynamic gene expression. *Genes Dev.*, 30(6):660–672.

D. Normanno, L. Boudarene, C. Dugast-Darzacq, J. Chen, C. Richter, F. Proux, O. Benichou, R. Voituriez, X. Darzacq & M. Dahan. 2015. Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat Commun*, 6:7357.

D. O'Carroll, H. Scherthan, A. H. Peters, S. Opravil, A. R. Haynes, G. Laible, S. Rea, M. Schmid, A. Lebersorger, M. Jerratsch, L. Sattler, M. G. Mattei, P. Denny, S. D. Brown, D. Schweizer & T. Jenuwein. 2000. Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. *Mol. Cell. Biol.*, 20(24):9423–9433.

H. Ogawa, K. Ishiguro, S. Gaubatz, D. M. Livingston & Y. Nakatani. 2002. A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. *Science*, 296 (5570):1132–1136.

P. Orth, D. Schnappinger, W. Hillen, W. Saenger & W. Hinrichs. 2000. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat. Struct. Biol.*, 7 (3):215–219.

N. Otsu. 1979. A threshold selection method from gray level histograms. *IEEE Trans. Systems, Man and Cybernetics*, 9:62–66.

M. Pal, A. S. Ponticelli & D. S. Luse. 2005. The role of the transcription bubble and TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell*, 19(1):101–110.

A. Papantonis & P. R. Cook. 2013. Transcription factories: genome organization and gene regulation. *Chem. Rev.*, 113(11):8683–8705.

H. Park, N. Y. Kim, S. Lee, N. Kim, J. Kim & W. D. Heo. 2017. Optogenetic protein clustering through fluorescent protein tagging and extension of CRY2. *Nat Commun*, 8(1):30.

S. H. Park, S. E. Yu, Y. G. Chai & Y. K. Jang. 2014. CDK2-dependent phosphorylation of Suv39H1 is involved in control of heterochromatin replication during cell cycle progression. *Nucleic Acids Res.*, 42(10):6196–6207.

M. C. Patel, M. Debrosse, M. Smith, A. Dey, W. Huynh, N. Sarai, T. D. Heightman, T. Tamura & K. Ozato. 2013. BRD4 coordinates recruitment of pause release factor P-TEFb and the pausing complex NELF/DSIF to regulate transcription elongation of interferon-stimulated genes. *Mol. Cell. Biol.*, 33(12):2497–2507.

G. P. Pathak, D. Strickland, J. D. Vrana & C. L. Tucker. 2014. Benchmarking of optical dimerizer systems. *ACS Synth Biol*, 3(11):832–838.

G. P. Pathak, J. I. Spiltoir, C. Hoglund, L. R. Polstein, S. Heine-Koskinen, C. A. Gersbach, J. Rossi & C. L. Tucker. 2017. Bidirectional approaches for optogenetic regulation of gene expression in mammalian cells using Arabidopsis cryptochrome 2. *Nucleic Acids Res.*

A. H. Peters, D. O'Carroll, H. Scherthan, K. Mechtler, S. Sauer, C. Schofer, K. Weipoltshammer, M. Pagani, M. Lachner, A. Kohlmaier, S. Opravil, M. Doyle, M. Sibilia & T. Jenuwein. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell*, 107(3):323–337.

A. H. Peters, S. Kubicek, K. Mechtler, R. J. O'Sullivan, A. A. Derijck, L. Perez-Burgos, A. Kohlmaier, S. Opravil, M. Tachibana, Y. Shinkai, J. H. Martens & T. Jenuwein. 2003. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell*, 12(6):1577–1589.

D. F. Pietras, K. L. Bennett, L. D. Siracusa, M. Woodworth-Gutai, V. M. Chapman, K. W. Gross, C. Kane-Haas & N. D. Hastie. 1983. Construction of a small Mus musculus repetitive DNA library: identification of a new satellite sequence in Mus musculus. *Nucleic Acids Res.*, 11(20):6965–6983.

I. Pinheiro, R. Margueron, N. Shukeir, M. Eisold, C. Fritzsch, F. M. Richter, G. Mittler, C. Genoud, S. Goyama, M. Kurokawa, J. Son, D. Reinberg, M. Lachner & T. Jenuwein. 2012. Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell*, 150(5):948–960.

Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar & R Core Team, 2017. *nlme: Linear and Nonlinear Mixed Effects Models.* URL `https://CRAN.R-project.org/package=nlme`. R package version 3.1-131.

L. C. Platanias. 2005. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat. Rev. Immunol.*, 5(5):375–386.

L. R. Polstein & C. A. Gersbach. 2015. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.*, 11(3):198–200.

L. R. Polstein, P. Perez-Pinera, D. D. Kocak, C. M. Vockley, P. Bledsoe, L. Song, A. Safi, G. E. Crawford, T. E. Reddy & C. A. Gersbach. 2015. Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. *Genome Res.*, 25(8):1158–1169.

O. Porrua & D. Libri. 2015. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell Biol.*, 16(3):190–202.

A. Rademacher, F. Erdel, J. Trojanowski, S. Schumacher & K. Rippe. 2017. Real-time observation of light-controlled transcription in living cells. *bioRxiv*. doi: 10.1101/132050.

I. U. Rafalska-Metcalf, S. L. Powers, L. M. Joo, G. LeRoy & S. M. Janicki. 2010. Single cell analysis of transcriptional activation dynamics. *PLoS ONE*, 5(4):e10272.

F. Ramirez, D. P. Ryan, B. Gruning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dundar & T. Manke. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, 44(W1):W160–165.

J. L. Ramos, M. Martinez-Bueno, A. J. Molina-Henares, W. Teran, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan & R. Tobes. 2005. The TetR family of transcriptional repressors. *Microbiol. Mol. Biol. Rev.*, 69(2):326–356.

S. Rea, F. Eisenhaber, D. O'Carroll, B. D. Strahl, Z. W. Sun, M. Schmid, S. Opravil, K. Mechtler, C. P. Ponting, C. D. Allis & T. Jenuwein. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, 406(6796):593–599.

E. J. Richards & S. C. Elgin. 2002. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell*, 108(4):489–500.

C. C. Robinett, A. Straight, G. Li, C. Willhelm, G. Sudlow, A. Murray & A. S. Belmont. 1996. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J. Cell Biol.*, 135(6 Pt 2):1685–1700.

J. T. Robinson, H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander, G. Getz & J. P. Mesirov. 2011. Integrative genomics viewer. *Nat. Biotechnol.*, 29(1):24–26.

A. Roopra, R. Qazi, B. Schoenike, T. J. Daley & J. F. Morrison. 2004. Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. *Mol. Cell*, 14 (6):727–738.

F. Roudier, I. Ahmed, C. Berard, A. Sarazin, T. Mary-Huard, S. Cortijo, D. Bouyer, E. Caillieux, E. Duvernois-Berthet, L. Al-Shikhley, L. Giraut, B. Despres, S. Drevensek, F. Barneche, S. Derozier, V. Brunaud, S. Aubourg, A. Schnittger, C. Bowler, M. L. Martin-Magniette, S. Robin, M. Caboche & V. Colot. 2011. Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J.*, 30(10):1928–1938.

S. Sainsbury, C. Bernecky & P. Cramer. Mar 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.*, 16(3):129–143.

V. Saint-Andre, E. Batsche, C. Rachez & C. Muchardt. 2011. Histone H3 lysine 9 trimethylation and HP1gamma favor inclusion of alternative exons. *Nat. Struct. Mol. Biol.*, 18(3):337–344.

A. Sakaguchi, D. Karachentsev, M. Seth-Pasricha, M. Druzhinina & R. Steward. 2008. Functional characterization of the Drosophila Hmt4-20/Suv4-20 histone methyltransferase. *Genetics*, 179 (1):317–322.

N. Saksouk, T. K. Barth, C. Ziegler-Birling, N. Olova, A. Nowak, E. Rey, J. Mateos-Langerak, S. Urbach, W. Reik, M. E. Torres-Padilla, A. Imhof, J. Dejardin & E. Simboeck. 2014. Redundant mechanisms to form silent chromatin at pericentromeric regions rely on BEND3 and DNA methylation. *Mol. Cell*, 56(4):580–594.

N. Saksouk, E. Simboeck & J. Dejardin. 2015. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin*, 8:3.

A. Saunders, L. J. Core & J. T. Lis. 2006. Breaking barriers to transcription elongation. *Nat. Rev. Mol. Cell Biol.*, 7(8):557–567.

T. Schalch & F. A. Steiner. 2017. Structure of centromere chromatin: from nucleosome to chromosomal architecture. *Chromosoma*, 126(4):443–455.

J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak & A. Cardona. 2012. Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, 9(7):676–682.

L. Schmiedeberg, P. Skene, A. Deaton & A. Bird. 2009. A temporal threshold for formaldehyde crosslinking and fixation. *PLoS ONE*, 4(2):e4636.

C. A. Schneider, W. S. Rasband & K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, 9(7):671–675.

G. Schotta, M. Lachner, K. Sarma, A. Ebert, R. Sengupta, G. Reuter, D. Reinberg & T. Jenuwein. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.*, 18(11):1251–1262.

M. K. Schuhmacher, S. Kudithipudi, D. Kusevic, S. Weirich & A. Jeltsch. 2015. Activity and specificity of the human SUV39H2 protein lysine methyltransferase. *Biochim. Biophys. Acta*, 1849(1):55–63.

D. C. Schultz, K. Ayyanathan, D. Negorev, G. G. Maul & F. J. Rauscher. 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.*, 16 (8):919–932.

Y. Shav-Tal, R. H. Singer & X. Darzacq. 2004. Imaging gene expression in single living cells. *Nat. Rev. Mol. Cell Biol.*, 5(10):855–861.

D. Shechter, H. L. Dormann, C. D. Allis & S. B. Hake. 2007. Extraction, purification and analysis of histones. *Nat Protoc*, 2(6):1445–1457.

C. E. Sing, M. Olvera de la Cruz & J. F. Marko. 2014. Multiple-binding-site mechanism explains concentration-dependent unbinding rates of DNA-binding proteins. *Nucleic Acids Res.*, 42(6): 3783–3791.

M. Smolle & J. L. Workman. 2013. Transcription-associated histone modifications and cryptic transcription. *Biochim. Biophys. Acta*, 1829(1):84–97.

D. Spensberger & R. Delwel. 2008. A novel interaction between the proto-oncogene Evi1 and histone methyltransferases, SUV39H1 and G9a. *FEBS Lett.*, 582(18):2761–2767.

B. L. Sprague, R. L. Pego, D. A. Stavreva & J. G. McNally. 2004. Analysis of binding reactions by fluorescence recovery after photobleaching. *Biophys. J.*, 86(6):3473–3495.

S. L. Squazzo, H. O'Geen, V. M. Komashko, S. R. Krig, V. X. Jin, S. W. Jang, R. Margueron, D. Reinberg, R. Green & P. J. Farnham. 2006. Suz12 binds to silenced regions of the genome in a cell-type-specific manner. *Genome Res.*, 16(7):890–900.

T. J. Stasevich, Y. Hayashi-Takanaka, Y. Sato, K. Maehara, Y. Ohkawa, K. Sakata-Sogawa, M. Tokunaga, T. Nagase, N. Nozaki, J. G. McNally & H. Kimura. 2014. Regulation of RNA polymerase II activation by histone acetylation in single living cells. *Nature*, 516(7530): 272–275.

B. D. Strahl & C. D. Allis. 2000. The language of covalent histone modifications. *Nature*, 403 (6765):41–45.

D. Strickland, Y. Lin, E. Wagner, C. M. Hope, J. Zayner, C. Antoniou, T. R. Sosnick, E. L. Weiss & M. Glotzer. 2012. TULIPs: tunable, light-controlled interacting protein tags for cell biology. *Nat. Methods*, 9(4):379–384.

Z. Su, M. D. Boersma, J. H. Lee, S. S. Oliver, S. Liu, B. A. Garcia & J. M. Denu. 2014. ChIP-less analysis of chromatin states. *Epigenetics Chromatin*, 7:7.

K. Sugaya, M. Vigneron & P. R. Cook. 2000. Mammalian cell lines expressing functional RNA polymerase II tagged with the green fluorescent protein. *J. Cell. Sci.*, 113 (Pt 15):2679–2683.

H. Sugisaki & S. Kanazawa. 1981. New restriction endonucleases from Flavobacterium okeanokoites (FokI) and Micrococcus luteus (MluI). *Gene*, 16(1-3):73–78.

L. Swint-Kruse & K. S. Matthews. 2009. Allostery in the LacI/GalR family: variations on a theme. *Curr. Opin. Microbiol.*, 12(2):129–137.

M. Tachibana, K. Sugimoto, M. Nozaki, J. Ueda, T. Ohta, M. Ohki, M. Fukuda, N. Takeda, H. Niida, H. Kato & Y. Shinkai. 2002. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.*, 16(14):1779–1791.

H. Tamaru & E. U. Selker. 2001. A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. *Nature*, 414(6861):277–283.

J. Tang, N. W. Cho, G. Cui, E. M. Manion, N. M. Shanbhag, M. V. Botuyan, G. Mer & R. A. Greenberg. 2013. Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination. *Nat. Struct. Mol. Biol.*, 20(3):317–325.

A. Taslimi, B. Zoltowski, J. G. Miranda, G. P. Pathak, R. M. Hughes & C. L. Tucker. 2016. Optimized second-generation CRY2-CIB dimerizers and photoactivatable Cre recombinase. *Nat. Chem. Biol.*, 12(6):425–430.

P. Therizols, R. S. Illingworth, C. Courilleau, S. Boyle, A. J. Wood & W. A. Bickmore. 2014. Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science*, 346(6214):1238–1242.

P. Thevenaz, U. E. Ruttimann & M. Unser. 1998. A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process*, 7(1):27–41.

H. Thorvaldsdottir, J. T. Robinson & J. P. Mesirov. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinformatics*, 14(2): 178–192.

D. Tischer & O. D. Weiner. 2014. Illuminating cell signalling with optogenetic tools. *Nat. Rev. Mol. Cell Biol.*, 15(8):551–558.

A. Tsumura, T. Hayakawa, Y. Kumaki, S. Takebayashi, M. Sakaue, C. Matsuoka, K. Shimotohno, F. Ishikawa, E. Li, H. R. Ueda, J. Nakayama & M. Okano. 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells*, 11(7):805–814.

T. Tumbar, G. Sudlow & A. S. Belmont. 1999. Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. *J. Cell Biol.*, 145(7):1341–1354.

S. Tyagi & F. R. Kramer. Mar 1996. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.*, 14(3):303–308.

M. Uesugi, O. Nyanguile, H. Lu, A. J. Levine & G. L. Verdine. 1997. Induced alpha helix in the VP16 activation domain upon binding to a human TAF. *Science*, 277(5330):1310–1313.

C. R. Vakoc, S. A. Mandat, B. A. Olenchock & G. A. Blobel. 2005. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell*, 19(3):381–391.

B. van Steensel. 2011. Chromatin: constructing the big picture. *EMBO J.*, 30(10):1885–1895.

L. Vandel, E. Nicolas, O. Vaute, R. Ferreira, S. Ait-Si-Ali & D. Trouche. 2001. Transcriptional repression by the retinoblastoma protein through the recruitment of a histone methyltransferase. *Mol. Cell. Biol.*, 21(19):6484–6494.

A. Vaquero, M. Scher, H. Erdjument-Bromage, P. Tempst, L. Serrano & D. Reinberg. 2007. SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature*, 450(7168):440–444.

O. Velazquez Camacho, C. Galan, K. Swist-Rosowska, R. Ching, M. Gamalinda, F. Karabiber, I. De La Rosa-Velazquez, B. Engist, B. Koschorz, N. Shukeir, M. Onishi-Seebacher, S. van de Nobelen & T. Jenuwein. 2017. Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *Elife*, 6.

S. Venkatesh & J. L. Workman. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.*, 16(3):178–189.

A. Villar-Garea, L. Israel & A. Imhof. 2008. Analysis of histone modifications by mass spectrometry. *Curr Protoc Protein Sci*, Chapter 14:Unit 14.10.

A. Vivante, E. Brozgol, I. Bronshtein & Y. Garini. 2017. Genome organization in the nucleus: From dynamic measurements to a functional model. *Methods*, 123:128–137.

M. J. Vogel, L. Guelen, E. de Wit, D. Peric-Hupkes, M. Loden, W. Talhout, M. Feenstra, B. Abbas, A. K. Classen & B. van Steensel. 2006. Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Res.*, 16(12):1493–1504.

F. Vollmuth & M. Geyer. 2010. Interaction of propionylated and butyrylated histone H3 lysine marks with Brd4 bromodomains. *Angew. Chem. Int. Ed. Engl.*, 49(38):6768–6772.

G. P. Wagner, K. Kin & V. J. Lynch. 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.*, 131(4):281–285.

D. Wang, J. Zhou, X. Liu, D. Lu, C. Shen, Y. Du, F. Z. Wei, B. Song, X. Lu, Y. Yu, L. Wang, Y. Zhao, H. Wang, Y. Yang, Y. Akiyama, H. Zhang & W. G. Zhu. 2013. Methylation of SUV39H1 by SET7/9 results in heterochromatin relaxation and genome instability. *Proc. Natl. Acad. Sci. U.S.A.*, 110(14):5516–5521.

H. Wang, W. An, R. Cao, L. Xia, H. Erdjument-Bromage, B. Chatton, P. Tempst, R. G. Roeder & Y. Zhang. 2003. mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. *Mol. Cell*, 12(2):475–487.

J. Wang, S. T. Lawry, A. L. Cohen & S. Jia. 2014. Chromosome boundary elements and regulation of heterochromatin spreading. *Cell. Mol. Life Sci.*, 71(24):4841–4852.

T. Wang, C. Xu, Y. Liu, K. Fan, Z. Li, X. Sun, H. Ouyang, X. Zhang, J. Zhang, Y. Li, F. Mackenzie, J. Min & X. Tu. 2012. Crystal structure of the human SUV39H1 chromodomain and its recognition of histone H3K9me2/3. *PLoS ONE*, 7(12):e52977.

W. Wang, M. Carey & J. D. Gralla. 1992. Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. *Science*, 255(5043):450–453.

A. K. Wong & J. B. Rattner. 1988. Sequence organization and cytological localization of the minor satellite of mouse. *Nucleic Acids Res.*, 16(24):11645–11661.

B. Wu, J. A. Chao & R. H. Singer. 2012. Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. *Biophys. J.*, 102(12):2936–2944.

Y. I. Wu, D. Frey, O. I. Lungu, A. Jaehrig, I. Schlichting, B. Kuhlman & K. M. Hahn. 2009. A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature*, 461 (7260):104–108.

C. Yang, E. Bolotin, T. Jiang, F. M. Sladek & E. Martinez. 2007. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene*, 389(1):52–65.

Z. Yang, J. H. Yik, R. Chen, N. He, M. K. Jang, K. Ozato & Q. Zhou. 2005. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol. Cell*, 19(4):535–545.

A. Yearim, S. Gelfman, R. Shayevitch, S. Melcer, O. Glaich, J. P. Mallm, M. Nissim-Rafinia, A. H. Cohen, K. Rippe, E. Meshorer & G. Ast. 2015. HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. *Cell Rep*, 10(7):1122–1134.

M. Zhang, F. Wang, S. Li, Y. Wang, Y. Bai & X. Xu. 2014. TALE: a tale of genome editing. *Prog. Biophys. Mol. Biol.*, 114(1):25–32.

V Appendix

V.1 Supplementary Tables

Table S1: Specification of selected modified histone peptides.

peptide location	peptide sequence	name	Mod1	Mod2	Mod3	Mod4	N terminus
A15	ARTKQTARKSTGGK APRKQ	H3 1-19	K9me3				free
C1			R8me2s	K9me3			
C7			R8me2a	K9me3			
C13			R8Citr	K9me3			
C23			K9me3	S10P			
C24			K9me3	T11P			
D1			K9me3	K14ac			
D10			R8me2s	K9me3	S10P		
D14			R8me2s	K9me3	T11P		
D18			R8me2a	K9me3	S10P		
D22			R8me2a	K9me3	T11P		
E2			R8me2a	K9me3	S10P	T11P	
E20			R2me2s	K4me1	K9me3		
E21			R2me2s	K4me2	K9me3		
E22			R2me2s	K4me3	K9me3		
E23			R2me2s	K4ac	K9me3		
F20			K4me1	R8me2s	K9me3		
F21			K4me2	R8me2s	K9me3		
F22			K4me3	R8me2s	K9me3		
F23			K4ac	R8me2s	K9me3		
F24			K4me1	R8me2a	K9me3		
G1			K4me2	R8me2a	K9me3		
G2			K4me3	R8me2a	K9me3		
G3			K4ac	R8me2a	K9me3		
H4			R2me2s	K4me1	R8me2s	K9me3	
H5			R2me2s	K4me2	R8me2s	K9me3	
H6			R2me2s	K4me3	R8me2s	K9me3	
H7			R2me2s	K4ac	R8me2s	K9me3	
H8			R2me2a	K4me1	R8me2s	K9me3	
H9			R2me2a	K4me2	R8me2s	K9me3	
H10			R2me2a	K4me3	R8me2s	K9me3	
H11			R2me2a	K4ac	R8me2s	K9me3	
I12			R2me2s	K4me1	R8me2a	K9me3	
I13			R2me2s	K4me2	R8me2a	K9me3	
I14			R2me2s	K4me3	R8me2a	K9me3	
I15			R2me2s	K4ac	R8me2a	K9me3	
I16			R2me2a	K4me1	R8me2a	K9me3	
I17			R2me2a	K4me2	R8me2a	K9me3	
I18			R2me2a	K4me3	R8me2a	K9me3	
I19			R2me2a	K4ac	R8me2a	K9me3	
K2	PRKQLATKAARKSA PATG	H3 16-35	K27me3				acetylated
K7			R26me2s	K27me3			
K12			R26me2a	K27me3			
K17			R26Citr	K27me3			
N3	GKGGAKRHRKVLR DNIQGIT	H4 11-30	K20me3				acetylated
N14			K16ac	K20me3			
N18			K12ac	K16ac	K20me3		
N22			R19me2a	K20me3			
O2			R19me2s	K20me3			
O6			R24me2a	K20me3			
O10			R24me2s	K20me3			

The positions listed here are the ones encircled in **Figure 14**.

Table S2: ChIP-Seq samples and basic specifications.

HiSeq ID	raw reads	IP	coverage	genotype	sample	background contribution
1487	126238704	input	2.31	wild type	wt1	0.62
1488	114686342	IgG	2.09			
1489	109344739	H3	1.99			
1490	113601462	H3K9me3	1.98			
1491	118669205	input	2.17		wt2	0.65
1492	129958570	IgG	2.36			
1493	82126699	H3	1.50			
1494	110258311	H3K9me3	1.90			
1301	115835411	input	2.11	SUV31H1 sn	E469-2	0.62
1302	122707297	IgG	2.21			
1303	121785813	H3	2.19			
1304	117870903	H3K9me3	1.9			
1305	151919655	input	2.78		E469-3_1	-
1307	105351171	H3	1.94			
1309	120396586	H3K9me3	2.10			
2200	117683344	input	2.14		E469-3_2	0.55
2201	90115380	IgG	1.67			
2202	79580116	H3	1.46			
2203	88390142	H3K9me3	1.59			
1309	116178090	input	2.15	SUV39H2 sn	E480-2	0.64
1310	122110572	IgG	2.25			
1311	118605873	H3	2.19			
1312	117322103	H3K9me3	2.14			
1313	120467720	input	2.22		E480-6	0.54
1314	96396546	IgG	1.80			
1315	91525182	H3	1.71			
1316	87757090	H3K9me3	1.63			
1499	114271340	input	2.10	SUV39H dn	dn2	0.60
1500	122879266	IgG	2.25			
1501	90311619	H3	1.67			
1502	91572300	H3K9me3	1.69			
2188	82647573	input	1.51		dn3	0.60
2189	74936236	IgG	1.38			
2190	89700084	H3	1.64			
2191	98388112	H3K9me3	1.81			
1624	82243808	input	1.56		E124-1	0.70
1625	95447144	IgG	1.79			
1626	87137900	H3	1.63			
1627	100862106	H3K9me3	1.89			
1628	114172526	input	2.12		E124-3	0.75
1629	110705642	IgG	2.04			
1630	82019725	H3	1.52			
1631	102203517	H3K9me3	1.89			
1809	124657774	input	2.17	wild type (mESC)	wt26, rep1	0.61
1810	99032939	IgG	1.79			
1811	118095580	H3	2.11			
1812	90097243	H3K9me3	1.42			
1813	109316591	input	1.88		wt26, rep2	0.50
1814	65533816	IgG	1.18			
1815	104532754	H3	1.88			
1816	79378787	H3K9me3	1.27			

The coverage was calculated from bedgraph files as shown in **Figure 47** using a Perl script. The background contribution corresponds to the parameter *b* yielded by MCORE (Molitor *et al.* 2017).

Table S3: H3K9me3 and expression changes of in wt iMFEs methylated genes.

					H3K9me3 fold change			RNA								
								H1sn			H2sn			dn		
gene ID	chr	start	end	RefSeq ID	H1sn	H2sn	dn	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}
D1Pas1	chr1	188791294	188794506	NM_033077	0.76	0.78	0.41	1.02	-0.01	NA	1.28	0.12	NA	1.02	-0.03	NA
Gm996	chr2	25430935	25435619	NM_001005424	0.83	0.33	0.36	4.32	0.46	0.823	3.11	0.14	NA	2.28	-0.23	NA
Sp5	chr2	70312979	70315783	NM_022435	0.77	0.06	0.17	4.52	0.23	0.921	2.53	-0.41	NA	2.02	-0.91	NA
Zfp120	chr2	149940142	149962414	NM_023266	0.58	0.58	0.78	188.12	-0.14	0.907	210.38	0.12	0.895	266.32	0.6	0.294
Gm14139	chr2	150007490	150019015	NM_001145863	0.83	0.63	0.82	35.7	-3.21	1.25e-05	36.25	-3.39	1.76e-06	47.69	-0.91	0.351
Zfp937	chr2	150043834	150070610	NM_001142411	0.71	0.63	1.06	200.03	-0.21	0.848	242.53	0.27	0.689	237.86	0.19	0.758
Gm14124	chr2	150083252	150096036	NM_001142410	0.86	0.65	0.9	25.6	-2.34	0.00551	31.19	-1.28	0.158	57.56	0.42	0.676
3300002I08Rik	chr2	150136672	150188501	NM_027017	0.74	0.6	0.95	12.37	-0.2	0.932	15.19	0.17	0.918	14.7	0.11	0.941
Zfp345	chr2	150296726	150310799	NM_001034900	0.78	0.7	0.82	53.73	-3.24	2.01e-06	67.34	-1.47	0.0455	73.44	-0.87	0.241
Blcap	chr2	157382097	157392097	NM_016916	0.92	0.51	0.78	246.88	-0.3	0.724	302.02	0.21	0.747	244.87	-0.3	0.714
Nnat	chr2	157385845	157388255	NM_010923	0.77	0.47	0.57	1.77	0.32	NA	2.53	0.59	NA	1.22	0.14	NA
1700028P15Rik	chr2	171782378	171788299	NR_040509	1.19	0.72	0.39	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Cbln4	chr2	171861835	171868966	NM_175631	0.99	0.47	0.19	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Gnas	chr2	174123359	174126303	NM_201617	1.06	0.5	0.43	1.52	0.22	NA	1.78	0.33	NA	3.41	1.12	NA
4930591A17Rik	chr2	179149640	179151585	NM_026596	1.29	0.72	0.22	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Rprl2	chr3	22150291	22150529	NR_004439	2.84	0.82	0.38	2234.83	-4	1.68e-06	4005.86	-0.36	0.428	4854.63	0.08	0.969
Sox2	chr3	34548926	34551382	NM_011443	0.46	0.08	0.11	1.02	-0.01	NA	27.53	2.78	0.004	1.22	0.14	NA
Flg2	chr3	93001194	93025298	NM_001013804	1.16	0.89	0.27	1.02	-0.01	NA	1.03	-0.01	NA	1.21	0.14	NA
Foxe1	chr4	46357065	46358181	NM_183298	1.56	0.15	0.13	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Zfp37	chr4	61850570	61869236	NM_001290353	0.77	0.67	0.92	76.01	0	1	129.24	1.15	0.0224	119.67	0.69	0.636
Zfp37	chr4	61850570	61869580	NM_001290351	0.77	0.67	0.83	76.51	0.02	0.993	129.24	1.15	0.0222	120.67	0.7	0.629
Foxd3	chr4	99322989	99325362	NM_010425	0.82	0.18	0.13	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Pou3f1	chr4	124334888	124337899	NM_011141	0.31	-0.01	0.15	1.28	-0.14	NA	1.8	0.09	NA	1.44	-0.12	NA
Hmgb4	chr4	127937455	127938139	NM_027036	1.3	0.67	0.38	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Gm13238	chr4	145427121	145427895	NR_033612	0.95	0.57	0.28	4.3	0.06	0.983	3.08	-0.4	NA	7.41	0.89	0.515
Zfp992	chr4	145728598	145749009	NM_001085522	1.23	0.69	0.58	183.57	-0.27	0.858	133.96	-1.63	0.000664	299.09	0.76	0.274
Gm13034	chr4	146034924	146036119	NR_030771	0.77	0.65	0.31	115.66	0.3	0.788	119.98	0.36	0.631	135.28	0.56	0.413
4933438K21Rik	chr4	146437474	146442545	NR_045446	1.29	0.82	0.35	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Znf41-ps	chr4	146514185	146556724	NR_040355	1.26	0.72	0.7	131.84	-0.15	0.924	101.05	-1.15	0.0361	204.11	0.74	0.303
Rps19-ps3	chr4	147195885	147196311	NR_033639	0.78	0.74	0.35	82.11	-0.19	0.899	82.75	-0.22	0.826	100.52	0.26	0.756
Gm833	chr4	152071648	152075031	NR_033138	1.27	0.87	0.16	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Ajap1	chr4	152747329	152856939	NM_001099299	1.1	0.61	0.12	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Arhgef16	chr4	153652578	153674004	NM_001112744	0.28	0.52	0.13	1.02	-0.01	NA	1.28	0.12	NA	1.61	0.42	NA
AI506816	chr5	23198078	23218485	NR_015554	1.51	0.82	0.58	19.66	-0.45	0.81	22.39	-0.19	0.896	13.66	-1.57	0.109
Nkx3-2	chr5	42152721	42155459	NM_007524	0.52	0.18	0.16	4.07	-0.81	0.614	3.86	-1.07	NA	4.87	-0.57	0.729
Gm15446	chr5	110362581	110370729	NR_040366	0.72	0.66	0.77	16.69	-1.01	0.447	19.7	-0.6	0.631	21.25	-0.35	0.746
Myo18b	chr5	113117895	113325382	NM_028901	0.3	0.64	0.23	1.02	-0.01	NA	1.03	-0.01	NA	1.81	0.54	NA
1700095B10Rik	chr5	113222311	113230721	NR_040676	0.34	0.71	0.28	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Grk3	chr5	113339497	113444558	NM_177078	0.3	0.53	0.22	3.77	0.64	NA	1.28	-0.14	NA	37.01	2.85	0.018
Crybb3	chr5	113504858	113510604	NM_001159650	0.32	0.22	0.11	1.02	-0.01	NA	2.53	0.63	NA	1.42	0.29	NA
2900026A02Rik	**chr5**	**113515342**	**113592333**	**NM_172884**	**0.16**	**0.1**	**0.09**	**147.53**	**5.6**	**8.19e-19**	**300.55**	**6.99**	**2.78e-35**	**264.91**	**6.54**	**2.19e-19**

					H3K9me3 fold change			RNA								
								H1sn			H2sn			dn		
gene ID	chr	start	end	RefSeq ID	H1sn	H2sn	dn	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}
Tmem211	chr5	113655928	113668283	NM_001033428	0.37	0.31	0.22	1.02	-0.01	NA	1.03	-0.01	NA	1.6	0.4	NA
NA	chr5	113672239	113709621	NR_027934	0.33	0.36	0.17	8.77	1.89	0.0852	6.28	1.56	0.176	27.71	2.57	NA
Sgsm1	chr5	113672239	113739806	NM_172718	0.36	0.38	0.18	8.52	1.86	0.0935	6.03	1.52	0.193	26.73	2.55	NA
Sgsm1	chr5	113686493	113739806	NM_001162965	0.37	0.39	0.18	7.53	1.5	0.223	4.3	0.89	0.521	17.33	1.99	NA
NA	chr5	113686493	113709621	NR_027936	0.32	0.38	0.16	7.78	1.53	0.207	4.55	0.94	0.487	18.31	2.02	NA
5930412G12Rik	chr5	129085004	129106562	NR_015517	1.36	0.64	0.16	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Fzd10	chr5	129106980	129109968	NM_175284	1.04	0.35	0.08	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Piwil1	chr5	129242120	129261349	NM_021311	1.35	0.82	0.16	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Zfp68	chr5	139044879	139060971	NM_013844	1.01	0.73	0.99	463.64	0.42	0.577	460.86	0.37	0.504	660.28	1.04	0.0124
Zfp12	chr5	143996841	144010513	NM_001289589	1.04	0.48	0.63	286.24	0.17	0.877	262.91	-0.1	0.898	368.85	0.65	0.165
Zfp316	chr5	144011373	144031701	NM_017467	1.09	0.5	0.52	203.68	0.12	0.933	181.84	-0.23	0.773	249.76	0.51	0.401
Smo	chr6	29685496	29711366	NM_176996	1.69	0.79	0.29	1258.55	0.78	0.112	1260.32	0.77	0.0395	1580.08	1.11	0.00242
Klf14	chr6	30906020	30908990	NM_001135093	1.02	0.81	0.73	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Repin1	chr6	48543881	48549081	NM_001079905	1.03	0.38	0.4	36.44	1.4	0.234	119.3	3.56	1.89e-10	41.19	1.63	0.16
AI854703	chr6	48578165	48583687	NR_027236	0.59	0.33	0.29	2.52	0.53	NA	2.28	0.51	NA	2.41	0.86	NA
Nap1l5	chr6	58855226	58857120	NM_021432	0.78	0.6	0.74	1.02	-0.01	NA	1.03	-0.01	NA	1.22	0.14	NA
Zfp606	chr7	13063653	13081584	NM_026112	0.82	0.64	0.8	198.69	0.3	0.804	187.24	0.12	0.893	240.7	0.64	0.216
Zfp110	chr7	13420158	13435933	NM_022981	0.86	0.66	0.53	246.92	-0.36	0.708	310.05	0.22	0.787	233.33	-0.47	0.552
Zfp180	chr7	24866915	24892727	NM_001045486	1.01	0.81	0.53	316.29	-0.16	0.897	295.77	-0.41	0.538	343.55	0.03	0.973
Zfp146	chr7	30946286	30954746	NM_011980	1.22	0.79	0.97	555.17	-0.31	0.748	582.39	-0.21	0.79	791.94	0.51	0.34
Scgb1b3	chr7	32160610	32161935	NM_001256073	1.86	0.71	0.34	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Scgb1b30	chr7	34880449	34885856	NM_001099330	0.59	0.53	0.46	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
AW146154	chr7	48734243	48755260	NM_001033530	0.72	0.59	0.73	32.07	0.21	0.912	45.51	0.91	0.271	62.84	1.43	0.0204
Dbx1	chr7	56886868	56892205	NM_001005232	0.88	0.33	0.3	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Klhl35	chr7	106614513	106622530	NM_028145	1.13	0.26	0.63	1.27	0.11	NA	1.03	-0.01	NA	1.02	-0.03	NA
Rbmxl2	chr7	114352958	114354430	NM_029660	0.92	0.49	0.19	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Zfp629	chr7	134750548	134757947	NM_177226	1.31	0.59	0.58	229.56	-0.17	0.901	228.61	-0.23	0.817	303.13	0.44	0.444
Nkx6-2	chr7	146765274	146768696	NR_027857	0.9	0.21	0.26	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Nkx6-2	chr7	146767118	146768696	NM_183248	0.97	0.15	0.22	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
6430531B16Rik	chr7	147158201	147164654	NM_001033465	0.78	0.39	0.47	1.27	0.11	NA	1.28	0.12	NA	1.22	0.14	NA
Cdkn1c	chr7	150644243	150646955	NM_009876	1.46	0.33	0.23	1.27	0.11	NA	3.03	0.74	NA	65.8	4.07	9.21e-05
Rprl3	chr8	3803124	3803361	NR_024198	0.49	0.2	0.17	3019.22	-6.29	1.83e-28	5369.85	-0.38	0.371	6724.22	0.1	0.96
Zfp958	chr8	4613169	4630231	NM_145591	0.7	0.58	0.9	171.78	-0.24	0.847	177.75	-0.19	0.828	202.57	0.15	0.859
Sox1	chr8	12395518	12399555	NM_009233	0.85	0.21	0.13	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
NA	chr8	22083372	22084427	NM_001012640	1.43	0.45	0.34	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Zfp930	chr8	71732944	71754438	NM_001013379	0.78	0.57	0.7	179.57	0	0.999	194.21	0.16	0.836	143.08	-0.49	0.74
Gm10033	chr8	71894923	71919443	NR_038043	0.75	0.73	0.91	63.42	-0.35	0.817	83.85	0.32	0.705	82.44	0.25	0.807
Isyna1	chr8	73118379	73121189	NM_023627	2.1	0.43	0.87	12.69	0.43	0.831	34.8	2.04	0.0107	17.76	0.91	0.496
Zfp599	chr9	22051873	22064339	NM_181419	0.97	0.62	0.74	110.28	0.18	0.885	98.92	-0.15	0.876	117.61	0.26	0.783
Rfpl4b	chr10	38540346	38541585	NM_001177783	0.33	0.66	0.62	1.27	-0.13	NA	1.28	-0.14	NA	1.62	0.04	NA
Polrmt	chr10	79198869	79209326	NM_172551	1.4	0.65	0.72	279.97	-0.28	0.76	302.41	-0.1	0.896	281.6	-0.25	0.759
NA	chr10	81167569	81184721	NM_001243067	0.84	0.7	0.73	77.56	0.93	0.295	69.14	0.72	0.352	87.27	1.07	0.125
NA	chr10	81205565	81233887	NM_199062	0.79	0.6	0.88	42.73	0.66	0.564	42.75	0.64	0.484	43.46	0.61	0.515
NA	chr10	81326678	81343831	NM_001243067	0.72	0.56	0.67	72.16	0.9	0.292	59.05	0.48	0.579	70.14	0.76	0.315

					H3K9me3 fold change			RNA								
								H1sn			H2sn			dn		
gene ID	chr	start	end	RefSeq ID	H1sn	H2sn	dn	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}
NA	chr10	81364655	81393225	NM_199062	0.8	0.61	0.89	299.72	0.63	0.335	306.32	0.65	0.164	278.74	0.38	0.588
Zfp873	chr10	81510871	81524331	NM_001024626	0.82	0.6	0.7	95.01	-0.71	0.42	110.39	-0.29	0.756	109.56	-0.26	0.739
AU041133	chr10	81590757	81615810	NM_001163064	0.81	0.69	0.87	122.43	0.3	0.798	126.51	0.35	0.635	133.58	0.43	0.515
Zfp938	chr10	81687601	81704020	NM_001105557	0.69	0.67	0.87	114.66	0.13	0.929	134.48	0.48	0.474	162.02	0.79	0.215
Washc4	chr10	83006685	83059218	NM_001033375	0.07	0.08	0.17	1464.6	0.5	0.393	1644.3	0.74	0.0615	1864.45	0.89	0.0704
Appl2	chr10	83062778	83111409	NM_145220	0.11	0.07	0.08	262.17	0.8	0.204	245.72	0.65	0.192	1150.26	2.57	0.00872
1500009L16Rik	chr10	83185609	83225507	NM_001145198	0.16	0.13	0.1	8.55	0.73	0.675	14.08	1.53	0.153	27.51	2.76	0.000335
Nuak1	chr10	83834063	83903216	NM_001004363	0.19	0.15	0.12	1050.72	0.42	0.706	1074.08	0.44	0.586	4458.75	2.46	0.00544
Platr7	chr10	84010161	84016083	NR_040658	0.2	0.2	0.14	1.02	-0.01	NA	1.03	-0.01	NA	35.95	2.6	NA
Tcp11l2	chr10	84039691	84077100	NM_146008	0.15	0.14	0.14	135.08	1.05	0.235	143.82	1.22	0.0418	390.62	2.31	0.0184
Polr3b	chr10	84085181	84189923	NM_027423	0.15	0.16	0.17	424.96	-0.2	0.841	466.3	0.01	0.988	1260.78	1.56	0.155
Rfx4	chr10	84218792	84369281	NM_001024918	0.26	0.24	0.2	1.27	0.11	NA	1.03	-0.01	NA	1.02	-0.03	NA
Rfx4	chr10	84300892	84369283	NM_027689	0.25	0.21	0.18	1.27	0.11	NA	1.03	-0.01	NA	1.02	-0.03	NA
4933416C03Rik	chr10	115548719	115550973	NM_001161855	0.24	0.2	0.13	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Kcnmb4	chr10	115854923	115910579	NM_021452	0.28	0.31	0.21	4.27	0.78	0.632	9.03	1.89	0.0761	46.18	3.69	0.000373
Kcnmb4os1	chr10	115855176	115858291	NR_028107	0.26	0.25	0.16	1.02	-0.01	NA	1.03	-0.01	NA	1.21	0.14	NA
1700030O20Rik	chr10	116160122	116166233	NR_045345	0.32	0.34	0.26	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Myrfl	chr10	116213600	116333935	NM_001033333	0.22	0.21	0.21	1.27	0.11	NA	1.53	0.23	NA	4.94	1.38	0.353
4933412E12Rik	chr10	116387617	116400335	NR_038025	0.36	0.3	0.24	50.6	-0.47	0.712	54.84	-0.3	0.774	109.26	0.96	0.445
D630029K05Rik	chr10	116401795	116406163	NR_027847	0.22	0.26	0.17	1.27	0.11	NA	1.53	0.23	NA	4.55	1.32	0.381
D630029K05Rik	chr10	116403329	116406163	NR_027846	0.26	0.24	0.17	1.27	0.11	NA	1.53	0.23	NA	4.55	1.32	0.381
Yeats4	chr10	116652196	116661563	NM_026570	0.15	0.09	0.07	272.93	0.49	0.529	298.12	0.67	0.171	1284.58	1.86	NA
9530003J23Rik	chr10	116670808	116675737	NM_029906	0.16	0.11	0.14	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Lyz2	chr10	116714596	116719328	NM_017372	0.19	0.19	0.14	15.09	-1.1	0.389	15.54	-1.12	0.308	20.78	-0.15	0.939
Lyz1	chr10	116724850	116729924	NM_013590	0.15	0.11	0.15	2.05	-0.44	NA	2.33	-0.35	NA	3.23	0.03	NA
Cpsf6	chr10	116781723	116814029	NM_001013391	0.13	0.08	0.16	1885.01	0.12	0.901	1922.62	0.13	0.837	3788.21	1.23	0.249
Cpm	chr10	117066555	117124408	NM_027468	0.12	0.14	0.1	3.03	0.47	NA	2.3	0.28	NA	35.77	2.45	NA
Mdm2	chr10	117125930	117147814	NM_001288586	0.08	0.09	0.08	1704.3	1.39	0.00402	2016.09	1.68	9.24e-05	4169.45	1.66	0.225
Slc35e3	chr10	117170733	117183414	NM_029875	0.13	0.12	0.08	231.64	-0.04	0.978	289.59	0.49	0.349	849.59	1.9	NA
Nup107	chr10	117187698	117229761	NM_134010	0.13	0.1	0.13	1134.86	-0.43	0.508	1140.44	-0.47	0.308	1273.56	-0.1	0.923
Gdf7	chr12	8304723	8308760	NM_013527	0.85	0.13	0.22	37.27	2.44	0.0138	5.28	0.9	0.516	5.59	1.31	0.336
5730507C01Rik	chr12	18521315	18542060	NM_001201330	1.11	0.63	0.61	42.24	-1.36	0.104	45.56	-1.16	0.115	67.84	0.09	0.937
1700030C10Rik	chr12	20810252	20821640	NR_015521	1.04	0.78	0.32	11.14	0.36	0.871	10.21	0.31	0.861	46.07	2.58	0.00289
2410018L13Rik	chr12	22990696	23046943	NR_015504	1.12	0.7	0.59	16.46	-1.13	0.363	21.22	-0.46	0.729	27.87	0.2	0.873
9030624G23Rik	chr12	24728066	24782134	NM_001256489	1.22	0.68	0.53	34.88	-0.63	0.647	41.49	-0.21	0.862	55.74	0.46	0.547
Insm2	chr12	56699903	56703004	NM_020287	1.44	0.13	0.27	1.27	0.11	NA	1.03	-0.01	NA	1.02	-0.03	NA
NA	chr12	57632923	57636093	NM_001146198	1.53	0.23	0.19	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
1700121N20Rik	chr12	107680861	107685876	NR_036593	1.19	0.79	0.22	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
Dio3	chr12	111517439	111519307	NM_172119	1.24	0.15	0.13	55.77	3.54	2.49e-05	24.28	3.08	0.000492	33.8	3.26	0.00297
2210016F16Rik	chr13	58481406	58486586	NM_027335	0.72	0.64	0.77	134.62	0.08	0.958	181.7	0.74	0.23	184.94	0.59	0.661
Zfp808	chr13	62231249	62275296	NM_001039239	0.66	0.6	0.8	105.13	-1.08	0.122	112.01	-0.93	0.118	121.61	-0.49	0.522
Gm3604	chr13	62469075	62484534	NM_001162910	1.03	0.68	0.93	43.32	-1	0.321	49.54	-0.62	0.512	49.3	-0.49	0.674
Zfp935	chr13	62554375	62568172	NM_178875	0.69	0.64	1.03	137.21	-0.11	0.936	146.09	0.02	0.987	148.3	0.05	0.942
Gm5141	chr13	62873559	62887168	NM_001256065	0.96	0.71	1.23	84.07	-0.9	0.228	84.62	-0.95	0.176	96.85	-0.38	0.688

					H3K9me3 fold change			RNA								
								H1sn			H2sn			dn		
gene ID	chr	start	end	RefSeq ID	H1sn	H2sn	dn	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}
Zfp369	chr13	65380161	65399103	NM_178364	0.76	0.78	0.67	235.25	-0.21	0.852	241.21	-0.2	0.781	237.7	-0.19	0.783
Gm10324	chr13	66214388	66223772	NM_001177832	1.09	0.63	0.69	1.52	-0.01	NA	1.78	0.1	NA	2.8	0.66	NA
2410141K09Rik	chr13	66519049	66542054	NM_183119	1.05	0.65	0.64	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
2210408I21Rik	chr13	77274796	77752940	NM_001145676	0.64	0.38	0.42	108.82	0.22	0.855	107.47	0.15	0.873	102.69	0.04	0.966
3110070M22Rik	chr13	120276063	120277191	NR_027974	1.09	0.5	0.15	3.27	0.53	NA	1.28	-0.14	NA	1.62	0.04	NA
Tmem267	chr13	120276845	120284312	NM_001039244	0.96	0.61	0.48	15.59	0.89	0.577	5.13	-0.47	0.779	32.13	2.53	5e-04
Sox21	chr14	118632455	118636252	NM_177753	1.22	0.13	0.28	1.02	-0.01	NA	1.03	-0.01	NA	2.22	0.69	NA
Gp1bb	chr16	18620411	18622496	NM_010327	1.04	0.24	0.4	69.72	-0.24	0.885	87.78	0.31	0.736	75.81	-0.05	0.97
Gp5	chr16	30307770	30310867	NM_008148	0.92	0.21	0.46	7.65	-0.58	0.76	11.99	0.25	0.88	7.75	-0.63	0.645
2700054A10Rik	chr17	13679886	13746960	NR_045437	1.52	0.7	0.25	5.3	0.18	0.942	8.58	0.85	0.532	6.05	0.4	0.815
Zfp960	chr17	17201066	17226595	NM_001163919	0.85	0.71	1.32	28.39	-0.72	0.594	32.12	-0.42	0.731	34.71	-0.18	0.879
Zfp97	chr17	17258346	17283842	NM_011765	0.81	0.63	1.34	186.1	-0.65	0.363	207.64	-0.35	0.581	205	-0.31	0.686
Zfp763	chr17	33153808	33170326	NM_028543	0.83	0.65	0.6	108.01	0.12	0.941	107.24	0.07	0.949	128.56	0.45	0.533
Olfr94	chr17	37333750	37334911	NM_001011518	0.91	0.31	0.27	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA
2610044O15Rik8	chr17	95213816	95234160	NM_153780	0.7	0.72	1.06	149.31	-0.19	0.869	167.15	0.07	0.934	162.03	0	0.996
Prob1	chr18	35810004	35814853	NM_001270646	0.83	0.15	0.12	42.63	-0.9	0.374	41.65	-1.05	0.238	29.83	-2.02	0.0414
Pcdhb22	chr18	37678006	37681073	NM_053147	0.79	0.73	0.44	449.36	0.7	0.259	331.62	-0.07	0.932	391.42	0.32	0.6
Slc25a2	chr18	37797031	37798377	NM_001159275	0.45	0.73	0.41	16.44	0.71	0.644	13.8	0.41	0.788	9.38	-0.33	0.828
Stxbp3-ps	chr19	9632095	9633738	NR_073559	0.62	0.51	0.29	47.66	0.8	0.456	40.39	0.47	0.663	44.69	0.64	0.405
Sox3	chrX	58144540	58146605	NM_009237	0.7	0.4	0.11	1.02	-0.01	NA	1.03	-0.01	NA	1.02	-0.03	NA

The genes listed are longer than 200 bp and showed an H3K9me3 level at least 3-fold higher than the genomic average in wt iMEFs. The 290026A02Rik gene depicted in **Fig. 22** is listed in bold.

Table S4: Common set of differentially expressed genes in SUV39H sn/dn iMEF cell lines compared to wt.

					RNA									H3K9me3			
					H1sn			H2sn			dn						
gene ID	chr	start	end	RefSeq ID	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	wt	H1sn	H2sn	dn
Sema3e	chr5	14025275	14256689	NM_011348	103.02	5.53	1.74e-18	204.03	6.20	2.85e-20	407.34	7.31	1.12e-23	0.59	0.61	0.72	1.27
Nlrp2	chr7	5250148	5302637	NM_177690	399.27	-5.37	2.05e-27	403.33	-6.38	1.91e-43	321.11	-7.20	1.72e-33	0.91	0.95	1.25	1.59
Arsj	chr3	126066769	126143292	NM_173451	195.28	5.88	8.70e-21	103.30	5.53	6.39e-19	316.44	7.11	2.71e-28	0.70	0.76	0.73	1.12
Pogk	chr1	168323741	168339959	NM_001142948	115.02	5.01	2.57e-12	254.78	6.34	1.14e-20	312.37	6.75	3.43e-18	0.72	0.71	0.62	0.64
2900026A02Rik	**chr5**	**113515342**	**113592333**	**NM_172884**	**147.53**	**5.60**	**8.19e-19**	**300.55**	**6.99**	**2.78e-35**	**264.91**	**6.54**	**2.19e-19**	**4.04**	**0.64**	**0.40**	**0.35**
Slc1a6	chr10	78243240	78277570	NM_009200	102.28	4.12	1.38e-07	208.55	6.08	4.69e-23	425.53	6.24	3.49e-14	0.99	0.93	0.95	0.99
Apod	chr16	31296277	31314682	NM_007470	413.9	-5.85	2.23e-32	416.96	-7.48	9.34e-42	337.28	-6.07	1.59e-29	0.54	0.81	0.73	0.54
Klhl29	chr12	5084273	5382488	NM_001164493	122.48	-4.25	3.03e-13	121.85	-5.63	4.18e-21	97.62	-6.06	3.14e-21	0.6	0.81	1.66	0.72
Fzd6	chr15	38837825	38869736	NM_001162494	53.34	3.37	1.93e-07	104.88	4.45	2.44e-15	321.67	6.05	8.82e-34	0.48	0.75	0.63	0.66
Fzd6	chr15	38837878	38869736	NM_008056	53.09	3.36	2.05e-07	105.38	4.45	2.07e-15	321.27	6.05	8.82e-34	0.48	0.75	0.63	0.67
Gfra2	chr14	71289936	71379645	NM_008115	303.61	-6.56	5.12e-35	308.74	-6.78	1.44e-37	248.84	-6.04	3.05e-24	0.92	0.84	1.37	0.82
Sytl2	chr7	97450864	97533680	NR_110348	73.27	3.43	8.08e-05	85.53	5.24	5.06e-16	198.99	5.95	6.74e-14	0.42	0.65	0.66	0.81
Kcnj15	chr16	95479164	95521865	NM_001039057	92.34	3.8	2.35e-07	151.13	5.17	1.9e-21	423.7	5.93	1.54e-15	0.72	0.92	0.9	0.74
Rab3c	chr13	110844395	111070414	NM_023852	89.24	-4.44	1.22e-13	92.32	-3.93	2e-10	71.83	-5.8	9.51e-18	0.63	0.76	0.85	1.37
Plin4	chr17	56240013	56249225	NM_020568	53.07	3.36	3.03e-05	86.86	4.29	3.25e-09	242.29	5.77	2.7e-15	0.3	0.52	0.32	0.25
Fam134b	chr15	25773052	25903451	NM_001034851	102.62	3.17	5.34e-05	85.44	3.62	3.69e-11	463.04	5.71	5.91e-22	0.44	0.68	0.71	0.69
Ms4a4d	chr19	11611338	11632956	NM_025658	247.86	-3.01	1.06e-08	233.86	-5.16	1.47e-25	186.61	-5.7	6.84e-21	0.62	0.92	1.02	1.49
Fam134b	chr15	25870456	25903451	NM_001277315	97.12	3.14	7.1e-05	80.19	3.58	1.4e-10	430.39	5.69	5.65e-22	0.41	0.58	0.63	0.81
Fam134b	chr15	25870580	25903451	NM_001277316	96.12	3.13	7.32e-05	79.94	3.58	1.47e-10	430.19	5.69	5.91e-22	0.41	0.58	0.63	0.81
Fam134b	chr15	25863306	25903451	NM_001277318	96.62	3.14	7.21e-05	78.94	3.56	2.3e-10	426.23	5.66	1.94e-21	0.4	0.62	0.61	0.76
Fam134b	chr15	25872433	25903451	NM_001277317	96.62	3.14	7.21e-05	78.69	3.56	2.42e-10	426.22	5.66	2e-21	0.4	0.59	0.63	0.81
St3gal1	chr15	66934436	67008444	NM_009177	1868.01	-3.05	9.79e-13	1837.18	-3.79	8.89e-27	1412.49	-5.63	3.52e-41	0.53	0.75	0.62	0.57
Rnf130	chr11	49838832	49918235	NM_021540	86.59	4.28	2.99e-13	108.88	4.73	2.5e-17	230.25	5.57	1.64e-18	0.42	0.57	0.69	1.09
Ccl11	chr11	81871333	81876457	NM_011330	74.19	-3.02	1.53e-07	69.81	-5.25	5.89e-15	56.29	-5.46	4.02e-15	0.44	0.61	0.67	1.15
Sp100	chr1	87546624	87606023	NM_013673	100.51	-3.82	1.61e-08	104.35	-3.63	1.89e-10	80.45	-5.38	4.45e-15	0.87	2.02	1.66	1.19
Cdh3	chr8	109034751	109080811	NM_007665	269.27	6.6	4.87e-31	849.28	8.34	2.54e-55	177.22	5.32	2.37e-09	0.74	0.85	0.61	0.59
Slc24a3	chr2	145068346	145467675	NM_053195	76.77	4.38	9.72e-11	77.78	4.71	2.46e-13	137.04	5.26	4.53e-12	0.88	0.71	0.96	1.08
Ece2	chr16	20629923	20645988	NM_177942	51.78	3.56	2.16e-07	123.8	5.05	5.77e-20	161.11	5.18	1.78e-14	0.39	0.47	0.28	0.5
Scara3	chr14	66538231	66572581	NM_172604	516.84	5.94	5.56e-29	310.38	5.65	4.08e-36	260.17	5.13	4.16e-18	0.93	0.87	0.85	0.69
Slamf8	chr1	174511507	174520699	NM_029084	55.16	-4.61	4.27e-11	56.23	-4.69	6.33e-12	45.27	-5.09	7.02e-12	0.88	0.67	0.77	0.66
Gbp2b	chr3	142257810	142282140	NM_010259	126.49	-5.1	1.16e-17	129.48	-4.82	5.2e-19	104.33	-5.01	4.67e-16	0.85	0.66	0.86	1.64
Pmaip1	chr18	66618257	66625212	NM_021451	134.78	5.69	2.67e-21	371.8	7.28	1.88e-39	115.51	5	1.62e-08	0.87	0.79	0.6	0.82
Il2rg	chrX	98459725	98463545	NM_013563	61.41	-4.63	3.78e-12	62.76	-4.69	7.49e-13	50.64	-5	9.34e-14	0.77	0.75	0.79	0.92
Spats2l	chr1	57831704	58005241	NM_144882	114.09	4.51	1.18e-16	58.13	3.6	5.9e-09	301.79	5	2.73e-09	0.58	0.69	0.78	0.89
Pax3	chr1	78097841	78193711	NM_001159520	39.04	-3.45	1.93e-06	38.64	-4.37	3.34e-09	31.02	-4.99	6.62e-10	0.63	0.86	0.83	1.03
Pax9	chr12	57796625	57812217	NM_011041	64.02	3.26	0.000269	38.53	4.18	4.96e-08	51.65	4.99	1.86e-10	0.93	0.31	0.48	0.33
Gng2	chr14	20691780	20795915	NM_001285911	181.27	4.54	9.39e-09	88.28	5.24	8.57e-14	55.28	4.95	1.39e-09	0.84	0.69	0.78	0.95
Gng2	chr14	20691780	20796471	NM_001285910	179.77	4.54	9.43e-09	86.78	5.2	2.03e-13	55.28	4.95	1.39e-09	0.84	0.69	0.78	0.95
Hoxd4	chr2	74560034	74567216	NM_010469	67.46	-3.9	5.3e-09	70.09	-3.49	1.39e-07	54.7	-4.95	1.55e-13	0.66	1.42	0.9	0.43
Itih2	chr2	10016217	10052310	NM_010582	701.1	5.61	1.3e-20	361.41	5.47	2.47e-35	449.24	4.94	2.52e-10	0.6	0.86	0.65	0.89
Spats2l	chr1	57902414	58005241	NM_001164566	104.09	4.37	3.17e-15	54.63	3.5	2.8e-08	280.4	4.93	4.72e-09	0.58	0.67	0.75	0.82

					RNA									H3K9me3			
					H1sn			H2sn			dn						
gene ID	chr	start	end	RefSeq ID	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	wt	H1sn	H2sn	dn
Dpt	chr1	166726862	166754397	NM_019759	1058.76	-5.57	4.75e-25	1061.77	-8.17	1.71e-57	873.96	-4.88	7.47e-13	0.58	0.8	0.8	0.9
Col22a1	chr15	71628905	71864657	NM_027174	53.89	-3.22	8.22e-07	53.7	-3.65	6.87e-08	42.24	-4.88	7.25e-12	0.64	0.75	1.51	1.26
Gbp4	chr5	105544785	105568605	NM_001256005	73.5	-3.75	5.83e-09	72.54	-5.31	1.78e-15	59.25	-4.86	1.64e-12	1.22	0.94	1.59	1.47
Gbp4	chr5	105544785	105568571	NM_008620	73.5	-3.75	5.83e-09	72.54	-5.31	1.78e-15	59.25	-4.86	1.64e-12	1.22	0.94	1.58	1.48
Slpi	chr2	164179805	164182243	NM_011414	177.09	-4.68	1.72e-22	179.68	-4.82	3.08e-20	145.5	-4.84	1.71e-17	1.9	3.01	1.46	0.59
Ifi27l2a	chr12	104672082	104681890	NM_001281830	48.69	-4.27	8.59e-09	49.23	-4.62	2.47e-10	39.9	-4.82	1.09e-09	0.14	0.62	0.67	0.55
Tgm2	chr2	157942140	157972128	NM_009373	153.28	5.01	7.99e-16	475.3	3.71	4.4e-05	251.22	4.8	2.62e-07	0.49	0.81	0.5	0.39
Ifi27l2b	chr12	104689107	104695433	NM_145449	45.21	-4.14	6.09e-09	48.41	-3.01	4.27e-05	37.02	-4.8	2.98e-10	0.48	0.87	0.55	1.1
Clec3b	chr9	123060063	123066550	NM_011606	55.69	-4.03	6.74e-09	56.13	-4.44	9.89e-11	45.4	-4.73	1.67e-10	0.66	1.52	0.84	0.68
Fgd4	chr16	16422070	16560242	NM_139232	32.02	3.79	9.68e-07	77.78	5.35	7.09e-16	67.72	4.64	5.94e-07	0.76	0.59	0.78	1.31
Col10a1	chr10	34109786	34116891	NM_009925	23.77	3.36	4.6e-05	101.78	5.43	4.59e-15	59.44	4.63	3.34e-07	0.59	0.61	0.5	0.41
Ifi27l2a	chr12	104680375	104681890	NM_029803	67.51	-4.51	1.53e-10	67.99	-5.07	4.93e-13	55.87	-4.6	1.91e-10	0.16	1	0.57	0.39
Robo1	chr16	72663393	73046345	NM_019413	300.07	4.41	2.32e-12	189.86	4.17	8.88e-14	288	4.59	3.95e-13	1.19	0.66	0.91	1
Ccdc3	chr2	5058821	5151912	NM_028804	63.97	-3.72	5.84e-09	64.96	-3.91	1.52e-10	51.81	-4.56	5.57e-11	0.81	1.03	0.88	0.89
Sec16b	chr1	159456300	159498555	NM_001159986	103.31	3.09	1.79e-07	109.48	3.27	1.44e-09	330.84	4.52	2.71e-12	0.56	0.57	0.54	0.67
Sec16b	chr1	159436926	159498555	NR_027641	105.07	3.06	2.41e-07	110.01	3.23	2.4e-09	338.65	4.5	3.29e-12	0.61	0.56	0.57	0.69
Eda	chrX	97170944	97596099	NM_001177942	79.02	4.63	2.69e-10	49.78	4.05	8.04e-07	49.41	4.49	7.44e-07	0.74	0.74	0.97	1.47
Fam65b	chr13	24730516	24825675	NM_029679	119.02	4.67	3.36e-12	101.03	4.93	7.36e-16	139.74	4.47	1.51e-06	0.59	0.64	0.65	0.68
Fam65b	chr13	24674057	24825675	NM_001286100	114.77	4.63	5.45e-12	98.78	4.9	1.53e-15	136.97	4.45	1.82e-06	0.66	0.72	0.69	0.71
Ddit4l	chr3	137286635	137291296	NM_030143	36.53	3.33	7.21e-05	111.55	5.2	1.11e-13	45.05	4.4	6.74e-08	0.81	0.59	0.84	0.71
Cacna1b	chr2	24459408	24618672	NM_001042528	53.02	4.07	1.69e-07	98.53	5.41	5.22e-15	62.35	4.34	9.36e-06	0.9	0.84	0.82	0.75
Slc6a17	chr3	107270465	107320936	NM_172271	177.27	6.24	2.4e-25	222.78	5.91	1.88e-16	53.97	4.33	6.71e-06	0.64	0.51	0.53	0.65
Bicd1	chr6	149357505	149511848	NM_009753	93.02	4.91	1.32e-14	95.28	5.15	6.59e-16	60.83	4.26	1.24e-06	0.89	0.73	0.88	1.4
Ifi209	chr1	175560989	175578059	NM_175026	26.2	-3.32	3.85e-05	26.6	-3.57	6.13e-06	21.42	-4.2	8.06e-07	0.69	0.68	0.94	1.89
NA	chr16	16435798	16504311	NM_139233	25.02	3.39	3.97e-05	55.03	4.85	3.71e-12	46.94	4.19	1.63e-05	0.72	0.57	0.74	1.44
Gngt2	chr11	95703981	95707045	NM_001038664	30.5	-3.14	8.49e-05	31.18	-3.28	1.34e-05	24.67	-4.18	5.07e-07	0.45	1.01	0.76	0.65
Gngt2	chr11	95703608	95707045	NM_023121	30.48	-3.13	9.26e-05	31.15	-3.27	1.45e-05	24.65	-4.18	5.79e-07	0.46	0.99	0.7	0.58
Sp140	chr1	87497277	87541611	NM_001013817	51.99	-3.73	3.77e-08	52.71	-3.88	1.1e-08	42.64	-4.17	7.61e-08	1.16	3.81	1.9	1.45
Gngt2	chr11	95704558	95707045	NM_001284397	30.23	-3.12	9.83e-05	30.9	-3.26	1.59e-05	24.45	-4.17	6.24e-07	0.46	1.05	0.76	0.65
Fhdc1	chr3	84246117	84284361	NM_001033301	62.12	3.11	5.62e-07	305.19	5.58	7.07e-37	129.23	4.08	1.44e-10	0.49	0.78	0.42	0.51
Fhdc1	chr3	84246117	84283413	NM_001205355	62.12	3.11	5.62e-07	306.19	5.58	6.26e-37	129.63	4.08	1.2e-10	0.51	0.81	0.44	0.54
NA	chr16	16453566	16560289	NM_139234	23.77	3.32	6.51e-05	50.28	4.72	2.48e-11	42.35	4.08	3.06e-05	0.72	0.6	0.78	1.21
Tnfsf15	chr4	63385636	63406147	NM_177371	26.52	3.35	6.46e-05	126.78	5.67	1.48e-16	33.87	4.06	1.63e-05	0.39	0.65	0.69	0.85
Has1	chr17	17980289	17992152	NM_008215	56.87	-3.81	2.84e-09	57.95	-3.98	2.46e-10	47.4	-4.06	1.87e-09	0.32	0.6	0.5	0.64
Ifitm1	chr7	148153974	148155726	NM_001112715	100.51	-3.33	3.23e-09	98.32	-4.33	1.14e-14	81.26	-3.97	2.52e-09	0.39	0.71	0.31	0.36
A530032D15Rik	chr1	85084713	85106428	NM_213615	26.51	-3.41	3.96e-05	27.06	-3.52	1.27e-05	22	-3.96	6.11e-06	1.23	4.09	1.94	1.47
Il1r2	chr1	40141612	40182070	NM_010555	25.9	-3.13	0.000106	25.77	-3.75	3.01e-06	21.15	-3.95	4.5e-06	0.5	0.81	0.77	0.78
C130026I21Rik	chr1	85242918	85267141	NM_001037909	40.11	-3.97	4.64e-08	40.75	-4.23	4.23e-09	33.88	-3.92	2.91e-07	1.18	4.09	1.9	1.37
Tbc1d2	chr4	46617261	46663071	NM_198664	38.55	3.45	4.63e-06	160.83	5.43	5.83e-17	46.75	3.91	2.18e-06	0.47	0.47	0.32	0.31
Ifitm1	chr7	148153327	148155726	NM_026820	101.28	-3.25	5.85e-09	98.6	-4.33	9.58e-15	81.88	-3.86	2.2e-08	0.37	0.5	0.21	0.4
Mndal	chr1	175787350	175810318	NM_001170853	144.23	-3.85	1.19e-15	145.03	-4.23	1.15e-18	119.81	-3.78	8.78e-08	0.56	0.88	0.68	1.34
Tram1l1	chr3	124023954	124026178	NM_146140	37.52	3.95	2.2e-07	94.28	5.47	5.44e-16	28.17	3.76	0.000142	0.84	0.71	0.76	0.49
Fam65b	chr13	24730516	24797870	NM_001080381	61.02	3.94	5.96e-08	55.28	4.22	4.73e-10	67.6	3.72	0.000261	0.59	0.58	0.57	0.68
Iqgap2	chr13	96397131	96661877	NM_027711	208.46	-4.53	1.69e-23	209.2	-5.32	3.52e-26	173.49	-3.7	9.9e-05	0.71	0.76	0.82	0.88

					RNA									H3K9me3			
					H1sn			H2sn			dn						
gene ID	chr	start	end	RefSeq ID	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	base mean	log2-fold change	p_{adj}	wt	H1sn	H2sn	dn
Eln	chr5	135178465	135223124	NM_007925	85.17	-3.29	1.86e-06	82.53	-5.22	1.03e-15	70.26	-3.7	1.24e-06	0.54	0.81	0.7	0.59
Armcx6	chrX	131282993	131285958	NM_001007578	81.82	4.23	1.12e-11	50.86	3.72	9.05e-09	47.53	3.66	1.39e-07	0.52	0.51	0.38	0.29
Fam65b	chr13	24706474	24797870	NM_178658	56.77	3.86	1.45e-07	53.03	4.15	1.25e-09	64.43	3.64	0.00041	0.62	0.59	0.58	0.68
Ccser1	chr6	61130318	62332857	NM_001164316	115.32	3.57	1.45e-08	87.86	3.24	1.58e-06	143.53	3.63	1.29e-05	0.83	0.6	1.02	1.62
Ano1	chr7	151774453	151924497	NM_001242349	32.52	3.38	1.49e-05	112.53	5.41	1.74e-18	55.3	3.57	0.000713	0.57	0.93	0.69	0.6
Xpnpep2	chrX	45461901	45490158	NM_133213	61.04	-3.62	5.11e-09	60.07	-4.82	6.29e-13	51.87	-3.57	1.96e-08	0.68	0.81	0.78	0.74
Arl4c	chr1	90594800	90598766	NM_177305	927.91	3.34	5.38e-23	1002.41	3.47	3.24e-29	1278.64	3.53	1.01e-13	0.22	0.34	0.05	0.1
Vnn1	chr10	23614493	23625149	NM_011704	34.02	3.62	7.71e-06	173.28	5.22	1.5e-11	15.71	3.36	0.000523	1.15	0.71	0.59	0.91
Arhgap22	chr14	34030008	34183122	NM_153800	54.84	3.37	1.33e-06	110.38	4.6	6.58e-16	50.19	3.32	3.04e-05	0.85	0.71	0.6	0.56
Lrrc32	chr7	105642731	105650340	NM_001113379	710.83	5.02	2.78e-34	836.66	5.31	2.06e-42	231.02	3.29	1.18e-12	0.35	0.75	0.2	0.31
Dio3	chr12	111517439	111519307	NM_172119	55.77	3.54	2.49e-05	24.28	3.08	0.000492	33.8	3.26	0.00297	3.87	4.79	0.57	0.51
Hs3st3b1	chr11	63698194	63735786	NM_018805	103.3	-3.67	8.68e-10	108.87	-3.14	8.33e-10	89.07	-3.25	4.86e-06	0.62	0.91	0.71	0.6
Fndc1	chr17	7931432	7997839	NM_001081416	94.27	5.1	2.5e-15	73.53	4.92	4.45e-13	219.07	3.23	0.00703	0.79	0.85	0.88	0.78
Thy1	chr9	43851466	43856662	NM_009382	789.2	-3.67	2.35e-05	800.43	-7.48	6.35e-59	644.87	-3.21	0.00888	0.58	0.79	0.46	0.34
Cers4	chr8	4493404	4526079	NM_026058	182.17	3.64	1.04e-09	116.27	3.19	9.02e-09	131.45	3.2	7.53e-07	1	0.61	0.85	0.74
Gramd1c	chr16	43980462	44016549	NM_153528	26.53	3.12	0.000112	35.3	3.7	8.84e-07	21.12	3.17	0.00043	0.66	0.58	0.59	0.85
Gramd1c	chr16	43980462	44028058	NM_001172107	26.53	3.12	0.000112	35.05	3.69	9.53e-07	21.12	3.17	0.00043	0.76	0.57	0.66	0.92
Deptor	chr15	54964990	55090828	NM_001037937	1023.04	-3.36	2.67e-25	1049.9	-3.24	1.59e-25	868.73	-3.11	2.42e-06	0.59	0.78	0.8	0.87
Deptor	chr15	54943871	55090828	NM_145470	1046.49	-3.36	1.71e-25	1073.75	-3.25	4.48e-26	889.32	-3.1	2.51e-06	0.57	0.78	0.75	0.81
Fat2	chr11	55064111	55125759	NM_001029988	130.17	-4.35	1.92e-17	134.58	-3.84	2.07e-11	112.11	-3.09	0.00212	0.75	1.11	1.01	0.89
Cercam	chr2	29725013	29738360	NM_207298	58.4	3.1	6.5e-07	124.49	4.31	3.98e-17	74.86	3.07	0.000635	0.29	0.6	0.33	0.34
Sytl4	chrX	130470923	130516322	NM_013757	20.02	3.1	0.000302	27.78	3.45	6.68e-05	31.78	3.06	0.00927	0.59	0.61	0.83	1.23
Dcxr	chr11	120586686	120588595	NM_026428	203.01	3.97	7.09e-18	211.15	3.91	6.09e-13	115.44	3.03	3.5e-08	0.11	0.47	0.27	0.43
Rhbdl3	chr11	80114413	80169488	NM_139228	127.74	-4.56	4.14e-15	137.5	-3.15	5.56e-08	117.26	-3.02	9.42e-08	0.41	0.9	0.65	0.59

Differentially expressed genes with a log2-fold change <-3 or >3 and an adjusted p-value <0.01 were selected for this list. The H3K9me3 value is the corrected enrichment of H3K9me3 above genomic average. The 290026A02Rik gene shown in **Fig. 22** is listed in bold.

V.2 Supplementary Figures

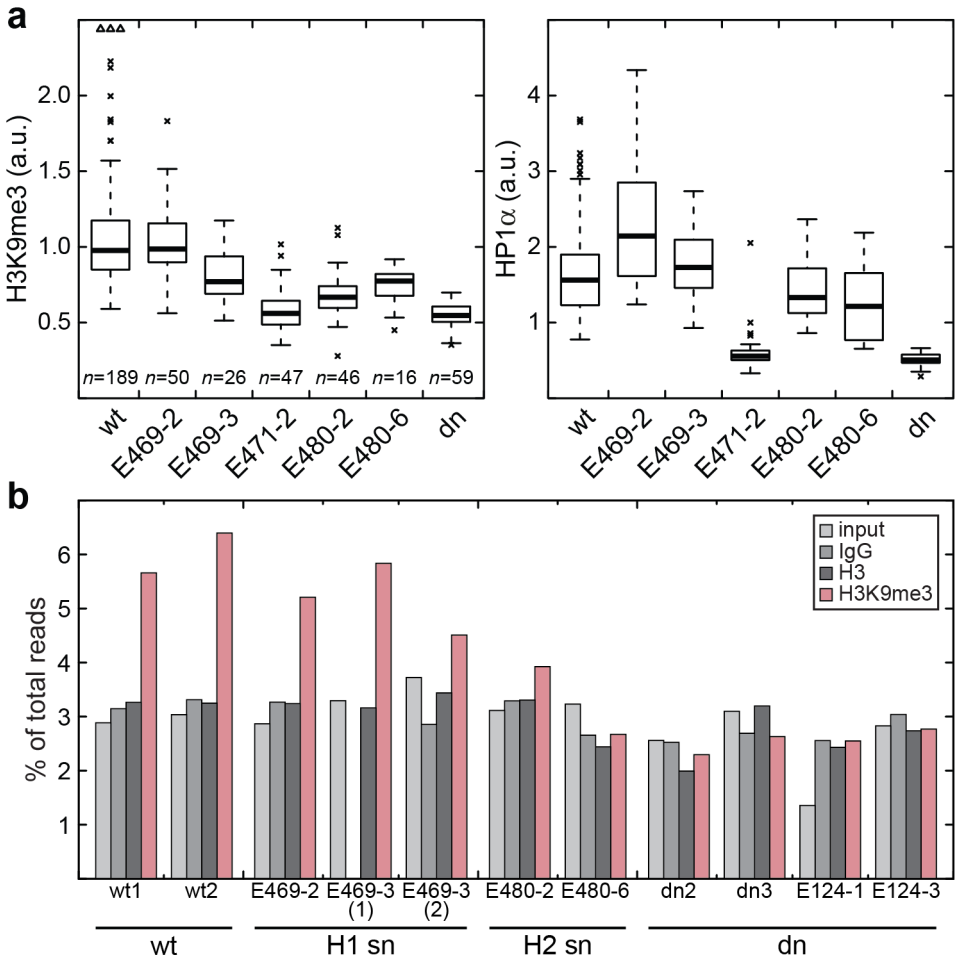

Figure S1: H3K9me3 at PCH in different SUV39H sn/dn iMEF cell lines. (**a**) Enrichment of H3K9me3 and HP1 at PCH as determined by CLSM after IF staining. Enrichments are normalized to chromatin density. (**b**) Percentage of ChIP-seq reads mapping to the major satellite consensus sequence (Lehnertz *et al.* 2003). The corresponding aggregated plots for cell lines with the same genotype are shown in **Fig 17c&d**.

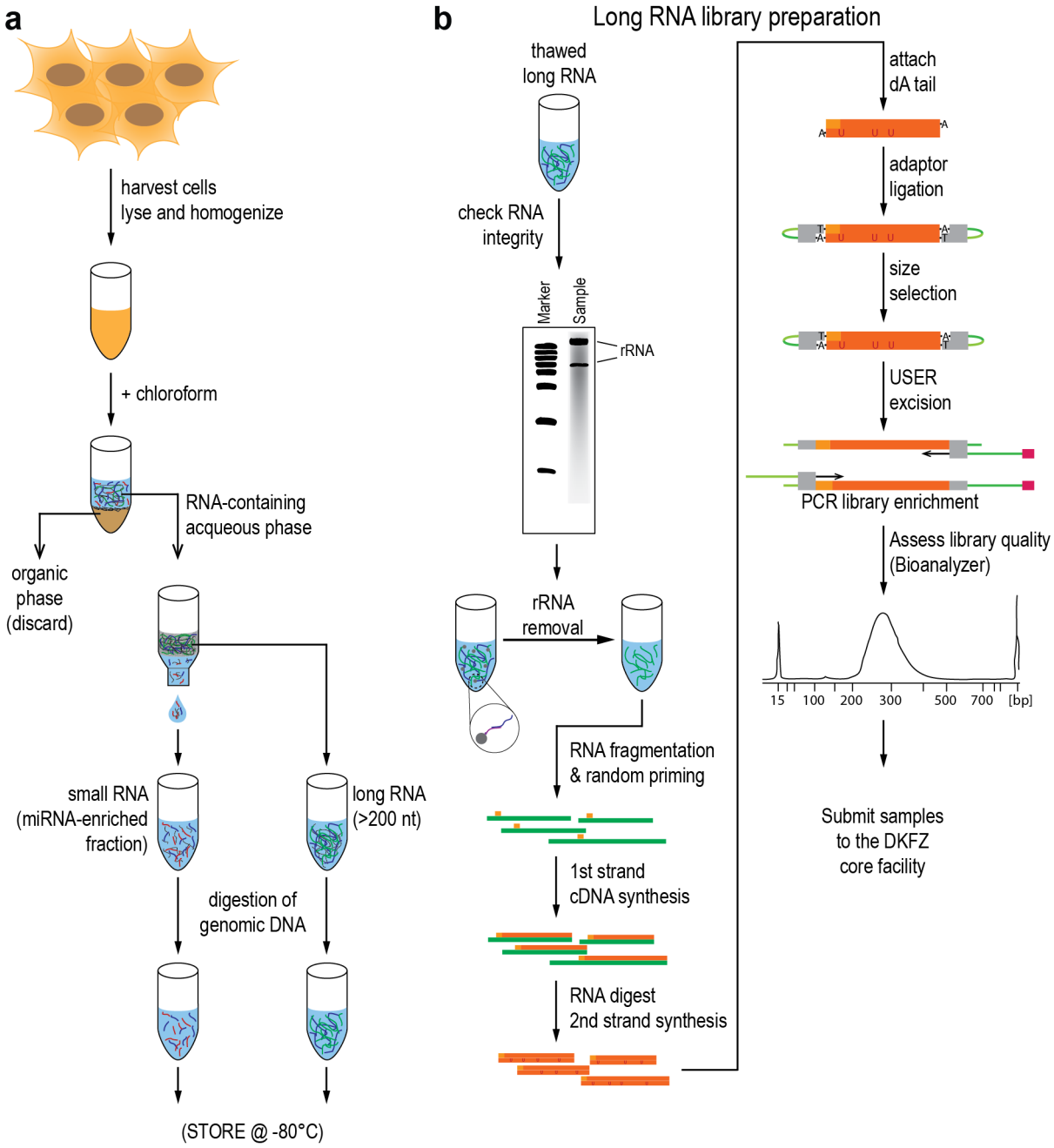

Figure S2: Stranded RNA-seq workflow. (**a**) RNA isolation with separation of small and long RNA. (**b**) Stranded long RNA library preparation. Details on the experimental steps are given in the Materials & Methods section.

VI Acknowledgements

Es ist ein lobenswerter Brauch: Wer was Gutes bekommt, der bedankt sich auch.

Wilhelm Busch

Ich danke besonders Karsten Rippe für seine Mühe und Geduld, mir die Wege der Wissenschaft näher zu bringen, sowie für seine großartige Motivation.

Desweiteren danke ich Harald Herrmann, Sylvia Erhardt und Benedikt Brors für das Interesse an meiner Arbeit und ihre Bereitschaft, als Zweitgutachter und Prüfer zu fungieren. Außerdem bin ich dankbar für die Begleitung und Diskussion der Arbeit im Rahmen meines TAC-Komitees.

Ich bedanke mich auch bei Gunnar Schotta für die Zusammenarbeit beim SUV39H-Teil meiner Arbeit.

Ein großer Dank gilt meinen Kollegen, insbesondere meinen Sitznachbarinnen und Freunden Katha und Jana, für die schöne Zeit bei der Arbeit und außerhalb bei zahlreichen Fußballabenden, Balkonkaffeepausen, Laborausflügen und Retreats. Besonders möchte ich mich bei Katharina und Fabian bedanken, die Mentoren für mich waren und immer ein offenes Ohr, eine helfende Hand und zur rechten Zeit auch eine kritische Bemerkung für mich hatten.

Bei meinen Freunden, vor allem vom Handball, bedanke ich mich für die Unterstützung. Kortine bin ich besonders dankbar für ihre Hilfe bei allen großen und kleinen Problemen.

Schließlich danke ich meiner Familie, insbesondere meinen Eltern, die von klein auf meine Neugier gefördert haben, und Franz, der meine größte Stütze ist und immer an meiner Seite steht.